UCLA
UCLA Electronic Theses and Dissertations

Title
Mechanisms and Modulation of Memory Processes within Subfields of the Human Hippocampus

Permalink
https://escholarship.org/uc/item/2bm9t9fz

Author
De Shetler, Natalie Grace

Publication Date
2016

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Mechanisms and Modulation of Memory Processes within Subfields of the Human Hippocampus

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Psychology

by

Natalie Grace De Shetler

2016

© Copyright by

Natalie Grace De Shetler

2016

ABSTRACT OF THE DISSERTATION

Mechanisms and Modulation of Memory Processes within Subfields of the Human Hippocampus

by

Natalie Grace De Shetler

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2016

Professor Jesse A. Rissman, Chair

Memory serves as a critical bridge between our present and our past, allowing previously formed associations to guide our present decision-making and inform our expectations about what is to come. Our ability to learn and retrieve certain types of associative knowledge, as well as to make projections about the world based on this knowledge, depends on the neural circuitry of the hippocampus, along with surrounding structures that comprise the medial temporal lobe (MTL) (Squire, 2009b). Declarative memory has been posited to rely on processes such as pattern completion and separation (Yassa & Stark, 2011), sequence encoding (MacDonald, Lepage, Eden, & Eichenbaum, 2011), and mismatch detection (Kumaran & Maguire, 2006). These processes enable memory to be flexibly encoded and retrieved so that it can inform all aspects of experience. This dissertation address three specific questions regarding how subfield functions work within a broader system of goals and motivations to facilitate flexible declarative memory. The project described in Chapter 3 focuses on the different degree to which discrimination and generalization responses, previously observed in

hippocampal subfields in response to close lure stimuli, is affected by task demands. Using high-resolution functional magnetic resonance imaging (hr-fMRI), we find diverging sensitivity to changes in stimuli and task demands in the responses within the left CA_1 and CA_3DG subfields. CA_3DG discriminates old visual stimuli from new stimuli, irrespective of the task-relevant context. In contrast, CA_1 responds to incorrectly paired repeats as it would to novels, while generalizes to different exemplars of the expected image. In Chapter 4, we use hr-fMRI investigate the impact of value on memory and show that items assigned a low reward value exhibit reduced encoding-retrieval neural similarity than either items assigned high reward value, or the similarity between items at encoding and unvalued items at retrieval. These results suggest that low value items may be processed in a less consistent manner within MTL regions than higher value items during encoding, possibly reflecting different encoding processes due to assigned values. Chapter 5 describes a set of behavioral studies aimed at evaluating whether mnemonic mismatches can lead to improved subsequent recollection of the item that violated memory based predictions. We show that when an item triggers a mismatch that item is recalled at greater rates than items that did not constitute a mismatch. Our results also show that this putative mismatch-induced encoding benefit is susceptible to disruption by numerous goal-related or motivational factors.

The dissertation of Natalie Grace De Shetler is approved.

Barbara Knowlton

Hugh T. Blair

Katherine Louise Narr

Nanthia A. Suthana

Jesse A. Rissman, Committee Chair

University of California, Los Angeles

2016

DEDICATION

To my Father, for being my inspiration, and to my Mother, for getting me there.

TABLE OF CONTENTS

LIST OF FIGURES ix

LIST OF TABLES xi

ACKNOWLEDGEMENTS xii

VITA xiii

CHAPTER 1
Introduction and Background 1

Introduction to the Dissertation 1
Hippocampal Structure and Imaging 3
Hippocampal Structure 3
Hippocampal Networks 7
High-Resolution Imaging 10
Proposed Roles of the Hippocampus and Subfields 17
Pattern Separation & Pattern Completion 17
Sequence Processing 24
Mismatch Detection 28

CHAPTER 2
Dissociable Profiles of Generalization/Discrimination in the Human Hippocampus during a Visual Associative Memory Retrieval Task 32

Hypotheses 34
Methods 34
Participants 34
Task Design 34
MRI Data Acquisition 36
ROI Creation, Registration and Pre-processing 36
fMRI Data Analysis 38
Results 39
Behavioral Results 39
Hippocampal Activity Tracks Memory Strength 40
Left Hippocampus Shows Associative Match Enhancement 41
Dissociable Profiles of Discrimination and Expectation Dependency 42
Differential Responses to Lure Probes 44
Differential Responses to Mispaired Probes 45
Discussion 45
Conclusion 48

CHAPTER 3
The Effects Of Reward Value On Item Representation In Human Hippocampus 50

Hypotheses 51
Methods 52
Participants 52
Experimental Design 52
Scanning Parameters 54
ROI Creation, Registration and Pre-Processing 55
Analyses 56
Results 60

Behavioral Data 60
Value Representation at Encoding 61
Encoding-Retrieval Similarity by Value 62
Effective Connectivity Between Reward Regions and MTL During Encoding 64
Value Representation at Retrieval 65
Discussion *66*
Conclusions *70*

CHAPTER 4
Stimulus Driven Shift between Memory Retrieval and Encoding: Encoding Benefits via Mismatch Detection **72**

Hypotheses 74
Experiment 1: Sequence Encoding and Reconstruction *75*
Design 75
Participants 77
Results 78
Discussion 78
Experiment 2: Sequence Reconstruction and Item Test *79*
Design 79
Participants 80
Results 80
Discussion 81
Experiment 3: Item Only Test *83*
Design 83
Participants 83
Results 83
Discussion 85
Experiment 4: Novel Items as Mismatch *88*
Design 88
Participants 89
Results 89
Discussion 92
Experiment 5: Effect of Memory Strength *94*
Design 94
Participants 95
Results 95
Discussion 98
Experiment 6: Effect of Prediction Strength: *99*
Design 100
Participants 101
Results 101
Discussion 102
Experiment 7: Response Prediction Mismatch *103*
Design 103
Participants 104
Results 105
Discussion 105
Experiment 8: Set Novelty *107*
Design 107
Participants 108
Results 109

Discussion 109
Meta-Analysis *110*
Repeated Images Recall 110
Novel Images Recall 111
General Discussion *112*
Conclusion *118*

CHAPTER 5
General Discussion **119**

References **124**

LIST OF FIGURES

Figure 1.1 4

Figure 1.2 8

Figure 1.3 21

Figure 2.1 36

Figure 2.2 37

Figure 2.3 39

Figure 2.4 42

Figure 2.5 43

Figure 2.6 44

Figure 3.1 53

Figure 3.2 55

Figure 3.3 57

Figure 3.4 60

Figure 3.5 64

Figure 4.1 78

Figure 4.2 81

Figure 4.3 82

Figure 4.4 85

Figure 4.5 87

Figure 4.6 90

Figure 4.7 97

Figure 4.8 101

Figure 4.9 **105**

Figure 4.10 **108**

Figure 4.11 **109**

Figure 4.12 **111**

LIST OF TABLES

Table 3.1 **61**

Table 3.2 **65**

Table 4.1 **75**

ACKNOWLEDGEMENTS

This research was supported by the following: UCLA startup funds and a Faculty Career Development Award provided to Jesse Rissman and a National Science Foundation grant (BCS-0848246) awarded to Barbara Knowlton.

I would like to thank my committee for all the help and support they have provided throughout this process. Foremost I thank my advisor Jesse Rissman who helped me develop as a research and scientist through his feedback and encouragement. I would also like to thank Barbara Knowlton, Nanthia Suthana, Tad Blair, and Katherine Narr for their guidance and feedback. I particularly want to thank Drs. Knowlton and Suthana for providing me the opportunity to work on the study that makes up chapter 3.

I am very grateful for the friends who supported and encouraged while at UCLA. Thanks to Joey for your unending willingness to help out whenever I needed it, and for our long talks with whiskey after work; to Nicco for our many debates and commiserations, and our shared love of spatial processing; to Tiffany for your willingness to help with anything and always getting my Futurama references; to Andrew for being there for the beginning and going through all the craziness that is grad school together; and to Julia for our many, many walks while discussing every aspect of grad school and beyond.

Finally I want to thank my family for everything they've done through the years to help me get to this point. I can never thank my Mother enough for everything she has done for me. Thank you to my brothers, Ben and Neil, for never letting me take myself too seriously.

VITA

2015 Candidate in Philosophy, Psychology – Cognitive Neuroscience
University of California, Los Angeles
Los Angeles, California

2012 Masters of Arts, Psychology – Cognitive Neuroscience
University of California, Los Angeles
Los Angeles, California

2009-2011 Staff Research Assistant, Semel Institute
University of California, Los Angeles
Los Angeles, California

2009 Bachelor of Arts, Psychology
University of California, Berkeley
Berkeley, California

2008-2009 Research Assistant, Helen Wills Neuroscience Institute
University of California, Berkeley
Berkeley, California

2007-2009 Research Assistant, Department of Psychology
University of California, Berkeley
Berkeley, California

PUBLICATIONS

De Shetler, N.G. & Rissman, J. (2016) Dissociable profiles of generalization/discrimination in the human hippocampus during an associative retrieval task. *Submitted.*

Ghahremani, D.G., Lee, B., Robertson, C.L., Tabibnia, G., Morgan, A.T., De Shetler, N., Brown, A.K., Monterosso, J., Aron, A.R., Mandelkern, M.A., Poldrack, R.A., London, E.D. (2012). Striatal dopamine D2/D3 receptors mediate response inhibition and related activity in fronto-striatal neural circuitry in humans. *Journal of Neuroscience*, *32*(21) 7316-7324.

PRESENTATIONS

De Shetler, N.G., Suthana, N., Nastarin, G., Yap, N., Knowlton, B.J., Rissman, J. (2015) Reward Effects on Encoding and Retrieval Representational Similarity in the Human Hippocampus. *Organization for Human Brain Mapping Annual Meeting.*

De Shetler, N.G., Clark, T.A., Rissman, J. (2013) Dissociable Profiles of Match/Mismatch Sensitivity in Hippocampus during Visual Associative Memory

Retrieval: A High-Resolution fMRI Study. *Society for Neuroscience Annual Meeting*.

Ghahremani, D.G. Lee, B., Tabibnia, G., Brown, A., De Shetler, N., Monterosso, J., Poldrack, R.A., London, E.D. (2011). Response inhibition reflects striatal dopamine receptor availability and BOLD signal. *Organization for Human Brain Mapping Annual Meeting*.

Morales, A., De Shetler, N., London, E.D. (2010) Changes in gray matter density after short-term abstinence from methamphetamine abuse. *Society for Neuroscience Annual Meeting*.

Mackey, A., Hill, S., De Shetler, N., Wendelken, C., Bunge, S. (2008) Does Reasoning Ability in Children Improve with Training? *Society for Neuroscience Annual Meeting*

HONORS and AFFILIATIONS

2014 UCLA Graduate Research Mentorship Award

2014 & 2012 UCLA Graduate Summer Research Mentorship Award

2012 UCLA Semel Institute NeuroImaging Training Program Fellowship, NIH 5T90DA022768-07

2011 UCLA Distinguished University Fellowship

2009 Psi Chi Psychology Honor Society

CHAPTER 1

Introduction and Background

Introduction to the Dissertation

Memory serves as a critical bridge between our present and our past, allowing previously formed associations to guide our present decision-making and inform our expectations about what is to come. Our ability to learn and retrieve certain types of associative knowledge, as well as to make projections about the world based on this knowledge, depends on the neural circuitry of the hippocampus, along with surrounding structures that comprise the medial temporal lobe (MTL) (Squire, 2009a). One of the defining characteristics of the hippocampus is its modularized structure. The hippocampus is made up of several sub-fields, including the dentate gyrus and the cornu ammonis regions (Duvernoy, Cattin, & Risold, 2013). Information flows through these sub-fields along a regular circuit, primarily along the tri-synaptic loop. Extensive rodent research has demonstrated that subfield functions are integral to understanding the numerous processes underlying the general phenomena of memory (Kesner, 2013c).

Declarative memory has been posited to rely on processes such as pattern completion and separation (Yassa & Stark, 2011), sequence encoding (MacDonald et al., 2011), and mismatch detection (Kumaran & Maguire, 2006). These processes enable memory to be flexibly encoded and retrieved so that it can inform all aspects of experience. These hippocampal-dependent faculties have resulted in a number of theories to explain the ultimate goal or function of memory (Eichenbaum, 2000; Mullally & Maguire, 2013). While a wealth of animal research provided valuable insights to

inform these theories (Allen & Fortin, 2013), results based on non-verbal species might have limited generalizability when applied to the rich experience of human declarative memory. Consequently, methodologies to investigate hippocampal subfield functions in healthy human subjects are of great importance and can provide further insights into the unique aspects of human memory.

In recent years, a number of hippocampal function experiments have utilized high-resolution fMRI (hr-fMRI) techniques (Carr, Rissman, & Wagner, 2010). Hr-fMRI allows for nearly twice the spatial resolutions as standard fMRI. With this resolution, experimenters can extract and isolate activities from each hippocampal subfield. A number of studies have used these techniques to great effect and have replicated findings from the animal literature in human subjects (Hassabis et al., 2009; Lacy, Yassa, Stark, Muftuler, & Stark, 2011). Furthermore, in combination with multivariate analyses, researchers have been able to decode specific memories, as well as to identify the content and/or context to which the hippocampus is sensitive (Bonnici et al., 2012; Chadwick, Bonnici, & Maguire, 2012; Chadwick, Hassabis, & Maguire, 2011; Chadwick, Hassabis, Weiskopf, & Maguire, 2010).

Below I will describe the current body of literature investigating hippocampal subfield functions, focusing particularly on those studies using hr-fMRI techniques on healthy subjects. I will then describe three studies that address several gaps in the literature. For each study, I will briefly describe the specific background and questions pertinent to the study before laying out the methodology used to address those questions. These studies address specific questions regarding how subfield functions

work within a broader system of goals and motivations to facilitate flexible declarative memory.

The study described in Chapter 2 focuses on the different degree to which discrimination and generalization responses, previously observed in hippocampal subfields in response to close lure stimuli, is affected by task demands. Chapter 3 also investigates how subfields' responses can be modulated by examining the effect of reward on item representation. Finally, Chapter 4 presents a series of experiments that investigate how mnemonic mismatch detection can allow for stimulus-driven switches between recall and encoding

Hippocampal Structure and Imaging

Hippocampal Structure

General organization. There are two main axes on which the hippocampus formation is organized: its long axis (anterior to posterior in humans and ventral to dorsal in rodents) and perpendicular to this axis, where the relative position of the subfields is the dominant feature (Duvernoy et al., 2013). It is the subfield organization that usually gets the most attention when describing the hippocampus's structure; however, this organization is not consistent along the long axis, which may give rise to different functions for the anterior and posterior hippocampus (Lacy & Stark, 2012; Poppenk, Evensmoen, Moscovitch, & Nadel, 2013; Poppenk & Moscovitch, 2011).

Cross-sectional organization. Two main structures make the subfields of the hippocampus, the dentate gyrus (DG), and the cornu ammonis (CA), which is commonly broken up into the CA_1, CA_2, CA_3, and the disputed CA_4 subfields. The DG and CA are wrapped around each other, forming two interlocking "C"s.

Between humans and rodents, the relative position of these interlocking “C”s are inverted, with the DG being inferior to the CA regions in the rodent, but superior in the human. Despite this flipped arrangement, the connections and flow of information between and within these regions is consistent. There are several paths of information through the hippocampus, but the one that is the most often described is the unidirectional "tri-synaptic loop". This loop describes the path that initiates in layer II of the entorhinal cortex (ERC) and first projects to the DG granule via the perforant path. Continuing the loop, the outputs of the DG form the mossy fibers, which synapse on the pyramidal cells of the CA_3 region, and the outputs of the CA_3 land on the CA_1 pyramidal neurons in the form of the Shaffer collaterals. Finally, CA_1 neurons project to the subiculum, which then flow back to the deep layers of the ERC (see Figure 1.1 for an illustration of these pathways). The flow of information through this loop, along with the characteristic of each region, has provided the framework for the prevailing theories of hippocampal function (Amaral, 1993; Duvernoy et al., 2013).

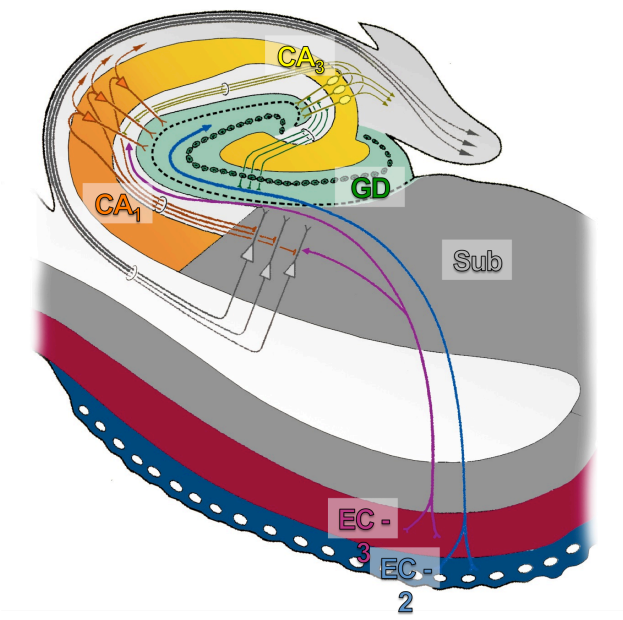


Figure 1.1: Schematic of the subfields of the human hippocampus. The pathways of the “tri-synaptic loop” are shown, as are direct projects from the ERC to CA_1. Adapted from Duvernoy et al.. (2013).

Despite the focus on this loop, there are a number of other circuits connecting regions. There are direction projections from layer III of the ERC to both CA_3 and CA_1, and CA_1 projects directly back to ERC. These

connections form two shorter loops from cortex to hippocampus and back (with and without CA_3) (Amaral, 1993; Duvernoy et al., 2013). These shorter loops seem to be quite sufficient for a number of processes since the removal of DG does not disrupt a number of memory tasks (Amaral, 1993). In addition to these interregional connections, both DG and CA_3 have a number of intraregional connections, CA_3 being particularly defined by its auto-associative network (Amaral & Witter, 1989).

Within the tri-synaptic loop, the nature of connections between the sub-regions is more complex that the loop describes. The perforant path is characterized by both a high degree of divergence, which each projection innervating many granule cells along the long axis of the hippocampus, and a high degree of convergence, with each granule cell receiving inputs from many ERC sources. The mossy fibers of the DG send more targeted, laminated, connection to CA_3, without much spreading along the long axis. The CA_3 to CA_1 connection shows divergence, sending signals along the length of the hippocampus (Amaral & Witter, 1989). Taken together, the numerous and diverse connections through the hippocampus allow for complex computations to be performed, which are essential for the demanding task of memory.

Long axis organization. While the most focus has been placed on the cross-sectional organization of the hippocampus, as mentioned above, the long axis of the hippocampus has its own organization. Two main features vary as one moves along the long axis, the relative proportion of the subfields and the major innervations from the cortex. In humans, the dentate gyrus makes up a larger proportion of the posterior hippocampus, whereas in the anterior hippocampus the CA regions hold the larger proportion (Duvernoy et al., 2013). As for connections, the anterior hippocampus

connects with numerous regions involved with emotion, reward and valence, including the amygdala, ventromedial prefrontal cortex (vmPFC), and the ventral tegmental area (VTA); whereas the posterior hippocampus is more connected with the parietal cortex and the cingulate. The entorhinal cortex provides the inputs to the hippocampus, and two main regions feed into the ERC, the perirhinal (PRC) and parahippocampal (PHC) cortices. The PRC forms most of the inputs that lead into the anterior hippocampus, while the PHC lead mostly to the posterior hippocampus (Kahn, Andrews-Hanna, Vincent, Snyder, & Buckner, 2008; Poppenk et al., 2013).

Poppenk, et al. (2013) have recently reviewed the functional differences of the long axis and have pointed to a number of different specializations. They found evidence for dissociations between anterior and posterior hippocampus for a number of processes, including: emotional vs. non-emotional memory, general cognition vs. spatial navigation, global vs. local spatial representation, gist vs. details (Poppenk et al., 2013). In an earlier paper Poppenk and Moscovich (2011) showed that the difference in relative volume and connections of the anterior and posterior hippocampus, rather than the overall hippocampus volume, could predict subjects' memory abilities (Poppenk & Moscovitch, 2011). This is further evidence for a functional distinction between the two segments.

A few studies have attempted to evaluate the functional connections between the PHC and PRC with subfields of the hippocampal formation using hr-fMRI. Lacy and Stark (2012) used a number of hr-fMRI data sets collected in their lab over several years to create a functional connectome of the MTL, including the hippocampal subfields. They found that the hippocampal formation's internal components were highly

interactive with each other, and the MTL cortexes were likewise strongly connected functionally; however, there was very little functional connection between the two regions (Lacy & Stark, 2012). Libby, Ekstrom, Ragland, & Ranganath (2012) also looked at the functional connectivity of the MTL, but broke up the region into anterior and posterior sections. They found increased connectivity between the anterior hippocampus and the PRC, while the posterior hippocampus was preferentially connected with the PHC (Libby, Ekstrom, Ragland, & Ranganath, 2012) consistent with the above noted long axis connections. The CA_1 and subiculum of the anterior hippocampus showed connectivity with the PRC, while the other subfields did not. In the posterior hippocampus, connectivity with the PHC was more uniform with all three subfields, though the subiculum showed the highest connectivity. The difference in findings between these two studies might reflect the different approaches to parcellating the MTL (including whether or not subdivide the hippocampus along its long axis), but it is clear that there is more work to be done looking at the functional connectivity of sub-regions. As Lacy & Stark note, functional connectivity does not always follow the structural connections.

Hippocampal Networks

Thus far, I have described the connections within the hippocampus formation, and the MTL. These connections reflect the unique position of the hippocampus as the final stage of a complex sensory processing schema, as described by Felleman and Van Essen (1991), where the incoming world is fully processed and then recorded as (or triggers recall of) a memory (Felleman & Van Essen, 1991). The hippocampus, however, has a number of connections outside of those emerging from sensory areas.

Two of these networks have received particular attention: the reward/valence system and the frontal lobes.

Hippocampus - reward network. Lisman & Grace (2005) describe a loop connecting the hippocampus to the nucleus accumbens (NAcc) and VTA and back. This loop is posited to originate in the CA_1 subregion of the hippocampus and its projections to the subiculum, see Figure 1.2. From there, subicular outputs reach the NAcc, which receives further inputs from the prefrontal cortex (PFC), as well as modulation from the VTA. NAcc then disinhibits the VTA, through inhibition of the inhibitory ventral pallidum. Finally, the VTA, which also receives salience information from the PFC and limbic system though pedunculopontine tegmentum (PPTg), triggers dopamine release back in the CA_1, as well as in the ERC (Lisman & Grace, 2005). The start and end of this loop has been shown to be most prominent in the body of the hippocampus (Kahn & Shohamy, 2013).

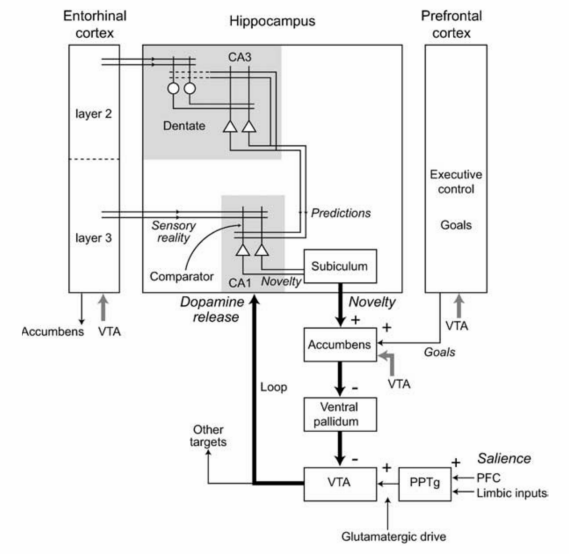


Figure 1.2: The Hippocampus-VTA Loop. From Lisman & Grace 2005.

This loop is posited to allow for both stimulus-driven and top-down goal-driven increases in memory encoding through enhanced long-term potentiation (LTP). Stimulus novelty is detected in CA_1, which is thought to act as a comparator, detecting when sensory reality, through the direct inputs from layer 3 of the ERC, mismatches with memory-based predictions emerging from the CA_3 pattern completion processes (Kumaran

& Maguire, 2007a) (see pattern separation/pattern completion section). This novelty signal is sent to the NAcc, where it can be modulated by PFC inputs that convey one's current goals. The combination of the bottom up and top down signals can trigger the influx of dopamine (Lisman & Grace, 2005). The validity and implication of this model will be further discussed in the mismatch detection section of this paper.

Frontal lobe and default mode connections. In addition to the PFC connections via the NAcc and VTA, the hippocampus has direct connections with the PFC (Simons & Spiers, 2003). As noted previously, the anterior (ventral in rodents) portion of the hippocampus is characterized as sending and receiving direct connections from the ventromedial prefrontal cortex. Preston & Eichenbaum (2013) recently reviewed a number of functions that have been ascribed to these connections. Prominent among them is that the PFC provides cognitive and strategic frameworks that influence which memories (or mnemonic details) are encoded and retrieved. The PFC can guide processing based on what is relevant to the moment while suppressing irrelevant information. Furthermore, the PFC helps to incorporate new information with existing schemas and thus facilitates encoding and memory consolidation (Preston & Eichenbaum, 2013; van Kesteren, Ruiter, Fernández, & Henson, 2012). The PFC may also support the integration of an ongoing event into a coherent episode (Ezzyat & Davachi, 2011).

Beyond the direct and anatomical connections described above, the hippocampus and the MTL, in general, are often considered to be components of the brain's “default mode network”. This network represents whole-brain functional connectivity between the medial prefrontal cortex, the MTL, cingulate cortex, precuneus,

and ventral lateral parietal cortex. This network has received a lot of attention over the last few years, but it is generally thought to support a mentalizing system, active when one is thinking about oneself, rather than in response to external stimuli (Andrews-Hanna, 2012). Reflecting on memories and planning for the future often comprise much of these internal ruminations, so the hippocampus is essential to this network (Buckner, 2010).

High-Resolution Imaging

Some work looking at hippocampal function uses patients with lesions to the hippocampus and MTL. These studies provided the first evidence tying the MTL to memory (Scoville & Milner, 1957; for reviews, see Squire, 2009a, 2009b). However, there are often only a limited number of subjects to support conclusions, such as the work done with the unique former London taxi driver with hippocampal damage, patient TT (Spiers & Maguire, 2007). Additionally, subjects can have other damage and/or some residual hippocampal tissue, such in the case of patient P01 who participated in research examining the consequences of hippocampal damage on one's ability to imagine future/counterfactual scenarios (Hassabis, Kumaran, Vann, & Maguire, 2007). Due factors such as these, there are strong limitations to what can be tested and concluded via human lesion investigations. In addition to lesion patients, patients with pharmacologically intractable epilepsy, who have electrodes implanted into their MTL in order to determine the source of their seizures, have been used to conduct studies more analogous to though performed with rats. These studies have, again, provided breakthrough findings, but are limited in manners similar to the lesion studies, namely

limited subject pools, and other sites of brain damage (e.g., from a lifetime of severe seizures).

Given the above-described limitations, fMRI has become the methodology of choice for studying the hippocampus in the human. The hippocampus is a relatively deep brain structure, so other noninvasive imaging methods that are best suited to recording from the surface of the brain, such as EEG are not very effective. FMRI allows robust recordings of hippocampal function, as well as fairly specific anatomical localization and has been widely used. Despite these strengths, the standard voxel resolution used in more fMRI studies, generally around 3 mm^3, does not allow for extraction of signal from the subfields of the hippocampus (Kirwan, Jones, Miller, & Stark, 2007). In order to investigate the subfield function, high-resolution fMRI (hr-fMRI) methods are required. In regards to imaging in the MTL, hr-fMRI is generally considered to imply the acquisition of voxels no larger than 2 mm^3, though there is a good degree of variance in voxel sizes across studies (Carr et al., 2010).

Functional MRI – Univariate Methods. Hr-fMRI presents a number of methodological challenges, as does the general use of fMRI to measure activity within the MTL. Initial fMRI studies of memory in the hippocampus encountered two major issues that often resulted in false negative results: first, the relatively small size of the hippocampus and its proximity to the sinus cavity (Stark & Squire, 2000) and, second, the elevated signal level of the hippocampus at resting baseline (Stark & Squire, 2001). The MTL sits above and around the sphenoid sinus, and the boundary between air and the denser bone and brain tissues results in magnetic field inhomogeneities, leading to signal drop-out in this region. Careful slice prescription, and selection of optimal TEs

(Milian et al., 2013), along with field unwarping techniques (such as FSL's FUGUE), can reduce this signal distortion, but it remains a problem.

The problem of the hippocampus's high rest baseline signal is much easier to address. As mentioned above, the hippocampus is an integral part of the default mode network. This network was first identified by its deactivation during many goal-directed cognitive tasks, and it is now commonly studied with a "resting-state" fMRI scan and functional connectivity analyses. During these scans, subjects are presented with no stimuli and are given no instruction as to any mental activity they should be engaging in other than staying awake. Subjects generally engage in self-mentalizing activities during these scans, which highly engages the hippocampus (Andrews-Hanna, 2012). In a conventional memory fMRI task, if task active periods are compared to a passive baseline (no task or stimuli presented), which, like the resting-state scans, engages the hippocampus, this severely reduces the ability to find an increased activation for the memory task, or even reduces the direction of the signal. The solution to this problem proposed by Stark & Squire (2001) is the use of an active baseline. Rather than comparing task period to a period of rest, it should be compared to a period of unrelated task. A simple task such as determining if a number is odd or even reduces the baseline activity of the hippocampus and increases the ability to detect results (Stark & Squire, 2001).

An additional consideration one has to take when conducting memory research in the hippocampus is the directionality of the BOLD signal. As mentioned above, the addition of an active baseline can reverse the direction of the percent signal change measured during the cognitive task of interest, but even when control for something like

baseline effects, there still remains a good deal of ambiguity in what is indicated by an increase or decrease in BOLD. In memory tasks, which often incorporate both encoding and retrieval components, it is hard to say if a decrease in signal reflects a real overall decrease of all hippocampus functioning, or a decrease of encoding in favor of retrieval, or a similar tradeoff (Reas & Brewer, 2013).

When it comes to hr-fMRI in the MTL these problems can be exaggerated as well as posing new issues. One of the challenges of hr-fMRI is the trade-off between resolution, acquisition time, and coverage, even with parallel imaging techniques that allow for more data collection in less time (Larkman & Nunes, 2007). In order to get the desired resolution, either longer TRs or a smaller field of views (FoV) are required. Researchers often avoid too much of a TR increase (which would diminish the temporal resolution of the acquired data), and rather choose a smaller FoV, which limits their ability to detect anything outside of their a priori interest areas. Along with the FoV tradeoffs, the signal to noise ratio (SNR) in fMRI is proportional to the size of the voxels being recorded from, so the reduction of voxel size needed to resolve subfields also reduces the signal extracted in these voxels. Averaging activity across many voxels within a region-of-interest (ROI) can help overcome some of the SNR loss (Kirwan et al., 2007; Poldrack, 2007), but hr-fMRI studies still tend to be underpowered and need many subjects to find reliable results (Carr et al., 2010).

Structural MRI. The use of anatomical ROIs comes with its own set of challenges. First, the hippocampal subfields need to be segmented for each subject. There are several MRI software programs that will segment the whole brain into various cortices and can be reliably used to make ROIs for the whole hippocampus, but

currently none are capable of further segmenting the hippocampus into subfields. Generally segmentation into subfields is done manually on a structural image acquired for such a purpose. There are a number of protocols for how the hippocampus should be segmented (Frankó, Insausti, Artacho-Pérula, Insausti, & Chavoix, 2014; Winterburn et al., 2013), and there is active debate about where to draw borders and what subfields should be segmented at all (the dentate gyrus is often not separated from the $CA_{3/4}$ regions, and rather a joint CA_3/DG ROI is used (Yassa & Stark, 2011). There have been some recent advances in automatic segmentation software (Yushkevich et al., 2010), which are growing in effectiveness.

The final step to using anatomical ROIs is registration. Due to the small size of the ROIs and their immediate proximity to each other there is a need for excellent registration between functional and anatomical images to ensure that one is extracting signal from the correct subfield (Kirwan et al., 2007). Additionally, if any group level analyses are desired, the cross-subject registration has to be particularly robust. Yassa and Stark (2009) compared a number of registration algorithms and concluded that the ROI-demons method yielded the greatest inter-subject registration overlap of ROIs (Yassa & Stark, 2009), though both authors' labs have since moved to using the Advanced Normalization Tools (ANTs) put out by Brian Avants (Avants et al., 2011).

Functional MRI – Multivariate approaches. Multivariate approaches have been gaining traction in the field of hr-fMRI work within the MTL. These approaches detect differences in the BOLD activity pattern associated with distinct mental states. This can be done across the whole brain or within a specific region. By comparing patterns between experimental conditions, rather than comparing averaged activity, subtle

differences in how the brain treats those conditions can be detected. Univariate methods of analysis require strong dissociations between various mental faculties to produce results, different conditions engaging a region to much greater or lesser extents, or engaging different brain areas entirely. Multivariate approaches do not require this dissociation; rather two conditions can both engage the same region, but with a different "neural signature". Thus multivariate findings provide valuable information about the type of information represented within individual brain regions, rather than simply indicating which regions are most active during a particular task (Rissman & Wagner, 2012).

Multi-Voxel Pattern Analysis. Multi-voxel pattern analysis (MVPA) uses machine-learning algorithms to detect the differences in BOLD activity patterns. One application of this method is decoding the content of memories. A pattern classifier can be trained to learn the difference in the brain's representation of unique stimuli, which is then used to test periods of recall and decode the likelihood that a given stimulus is being remembered (Chadwick et al., 2012). This method is effective since recall triggers the reactivation of the brain areas that processed the memory initially (Polyn, Natu, Cohen, & Norman, 2005). Decoding memories can be employed to determine which factors influence the fidelity of encoding and retrieval, such as task goals and competition between individual memories (Kuhl, Rissman, & Wagner, 2012).

MVPA has also been used to decode information from the hippocampus itself. One of the first MVPA studies using the hippocampus alone was able to decode a subject's position within a virtual environment (Hassabis et al., 2009). Individual episodic memories have also been decoded based on their distinct BOLD

representation within the hippocampus (Chadwick, Hassabis, Weiskopf, & Maguire, 2010). These studies show what types of detailed information is stored in the hippocampus. Univariate methods would never be able to provide these insights, as the mean activity in the hippocampus would be identical between different memories. It is only by looking at the activity patterns within the hippocampus that we can learn that not only is the hippocampus engaged in the recall of memories, but also some elements of the individual memories are distinctly represented there as well.

Detecting, or failing to detect, differences in the patterns that represent slightly different stimuli has been used to assess how distinct hippocampal subfields perform pattern separation and/or pattern completion (see pattern separation section for a detailed explanation of these terms). Bonnici (2012) demonstrated a combination of MVPA analysis and hr-fMRI to decode within the subfields of the hippocampus (Bonnici et al., 2012). This combination allows for very nuanced questions to be asked about what the hippocampal subfields are sensitive to, and what they can represent. Apart from asking question directly relating to the subfields, it is unclear if hr-fMRI can provide additional information for decoding; the increased number of features that a high-resolution image would provide might allow for more nuanced patterns to be detected or might simply result in more noise (Chadwick et al., 2012).

Representational Similarity Analysis. Another multivariate approach that has been used to detect what information is represented within the MTL is representational similarity analyses (RSA). Rather than using machine learning, RSA calculates the correlation between brain activity patterns evoked during the processing of different stimulus types or experimental conditions (Kriegeskorte, 2008). This correlation serves

as an estimate of how similarly that region represents stimulus information and/or cognitive processes associated with those respective conditions. As with MVPA, studies employing RSA have revealed coding of information related to the spatial location and temporal remoteness of individual episodic memories (Nielson, Smith, Sreekumar, Dennis, & Sederberg, 2015), as well as reward information (Wolosin, Zeithamova, & Preston, 2013). RSA has shown that increased neural similarity for items across successive encoding opportunities predicts subsequent recall (Xue et al., 2010). RSA that examines the similarity between stimuli presented at the time of encoding and again at retrieval, also known as Encoding Retrieval Similarity (ERS), has revealed that greater ERS estimates is indicative of improved recall performance (Ritchey, Wing, LaBar, & Cabeza, 2013), possibly as a results of greater reinstatement of the learning context.

Proposed Roles of the Hippocampus and Subfields

Pattern Separation & Pattern Completion

The topic that receives the most attention in work on hippocampal subfields is pattern separation and pattern completion. These two processes are the proposed solutions to two of the biggest challenges to any memory system: 1) how to efficiently encode similar events so that they can be dissociated later (e.g., preventing the memory of where one parked one's car yesterday from interfering with encoding and later recalling where it is parked today); 2) how to retrieve memories when the cues are degraded or slightly altered (e.g., remembering where one's car is even though the time of day has changed and the parking lot looks different).

Theoretical model. In memory terms, pattern separation refers to the taking of two similar items or events and creating representations that are maximally orthogonal to each other; pattern completion is the process of triggering recall from degraded stimuli (Hunsaker & Kesner, 2013). As seen in Figure 1.3a, both pattern completion and pattern separation operate on stimuli that are slightly different from each other, A & A', though pattern separation emphasizes the differences, while pattern completion minimizes them (Yassa & Stark, 2011).

An important note about these processes is that each refer only to computations done during one phase of memory: pattern separation occurs during encoding as pattern separation implies the building of a new representation, and pattern completion happens during retrieval, as completion implies triggering previously stored memories (Hunsaker & Kesner, 2013). In the literature, including some that I will review here, there is a tendency to invoke the terms pattern completion and separation to describe memory phases that each term is not appropriate for. Furthermore, failures to pattern separate are often treated as pattern completion, or visa versa. This can be especially confusing when we acknowledge that we switch flexibly and rapidly between encoding and retrieval in many, or most, situations. An example of an instance where the term pattern separation has been applied in a confusing manner can be seen in a behavioral study that concluded that recall could occur both with and without pattern separation. By presenting stimuli that were close lures to previously learned items, and asking subjects to determine if the stimuli was "old", "new", or "similar", the study showed that recall was just as likely without what was referred to as pattern separation (false alarming to foils, calling them "old") as it was with (recognizing items as "similar"). While the findings of

the study are sound, pattern separation might not be the correct term to refer to determining lures as "similar" rather than "old"; the notion of "recall to reject" is sometimes posited as the mechanism for identification of lures (Kim & Yassa, 2013).

Biological model. The processes of pattern separation and pattern completion are particularly of interest to those who evaluate subfield function since the cytoarchitecture of these fields, especially the dentate gyrus and CA_3, lend themselves well to the two processes. As described before, the DG receives many converging and diverging inputs from the ERC, but fires sparsely and sends sparse connection to area CA_3 (Amaral & Witter, 1989). Given this structure, the DG is well suited for pattern separation through sparse encoding (Kesner, 2013). The CA_3 is characterized by its highly recurrent connections, which are thought to perform pattern completion through this auto-associative network (Kesner, 2013a). The ERC→ DG → CA_3 stream seems well designed for completing both processes (Yassa & Stark, 2011). Additionally, back projections from CA_3 to the DG might also help with the efficacy of this network's ability to perform pattern separation by providing feedback from the results of pattern completion (Myers & Scharfman, 2011). The mechanisms of how pattern separation is performed in the DG are complicated, however, by recent findings that while young DG cells (the dentate gyrus being one of the few areas that exhibit adult neurogenesis) show pattern separation activity, older cells in the DG are more likely to perform pattern completion (Nakashiba et al., 2012).

Investigations into the topics of pattern separation and completion are very much ongoing, especially in rodent research, where, as noted above, direct recordings can be made from the specific sub-region in question. When conducting investigations in

humans, however, direct measures are not possible, and more general inferences are required. An often-used method of inference employed in human pattern separation research is to look for regions whose response properties adhere to a computationally defined manifestation of pattern completion and pattern separation. This definition focuses on the change in output of a region (via changes in BOLD signal in fMRI studies) from change in input (via parametric changes in stimuli relative to a previously learned baseline). Pattern separation is characterized by a large change in output for a small change in input while pattern completion is characterized by a small change in output to larger changes in input. A number of studies in rodents find large changes in DG activity in response to small changes in learned environments, which is generally called pattern separation (the sparse firing of DG granule neurons make recording from the DG more difficult than other subfields, but this finding has been seen in a few studies that have managed it) (Kesner, 2013a). The firing patterns of CA_3 recordings tend to be more robust (little change in output) to stimuli changes, indicative of pattern completion, when the change in stimuli is fairly small. In the presence of larger changes in stimuli, the firing pattern of this region is modulated to a much greater extent, reflecting a shift to pattern completion (Guzowski, Knierim, & Moser, 2004; Kesner, 2013b).

These two changes, change in input and change in output, when plotted against each other, form the graph seen in Figure 1.3b. The curve representing DG firing forms a positive power function, or saturation curve; the firing of CA_3 forms an S-shape curve. The validation line of equivalent change in output for change in input is used to demark what can be considered pattern completion (below the line) and pattern separation

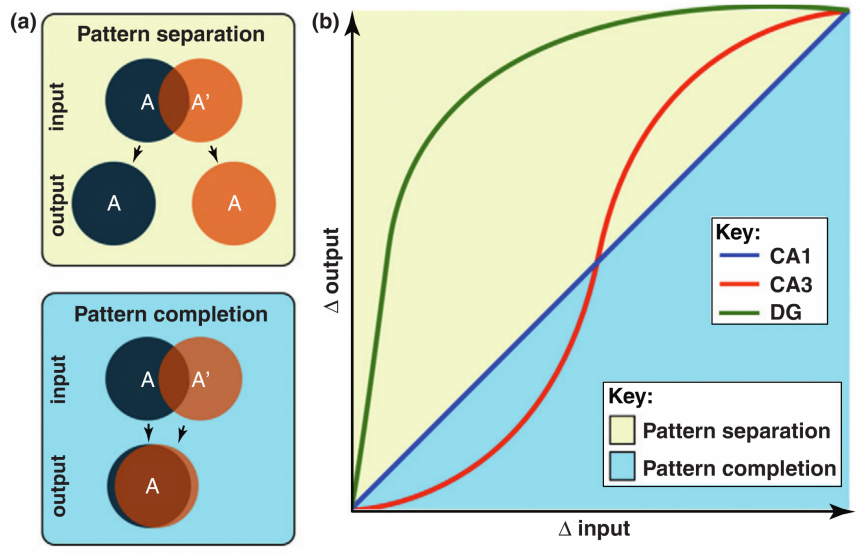


Figure 1.3: Theoretical Models of Pattern Separation and Completion. From Yassa & Stark 2011.

(above the line) (Yassa & Stark, 2011). This linear change line is often labeled as the CA_1 firing pattern, which generally, though not always, is thought to play a role outside of pattern completion/separation (see *Mismatch Detection* for more focused discussion of CA_1 function). It is important to remember that pattern completion reflects recall and separation reflects encoding (Hunsaker & Kesner, 2013), so the switch between the two in response to increased change in input could be described as a switch to encoding once a stimulus is sufficiently novel.

Evidence for pattern completion and pattern separation in humans. ***Standard resolution fMRI.*** Montley and Kirwan (2012) used this computational definition of pattern completion and pattern separation to demonstrate dissociable sensitivity to types of pattern separation between the right and left hippocampus. They presented stimuli that were either a previously learned item, a new item, or an old item parametrically rotated to various degrees and asked subjects to either determine what category the item was ("toy" or "not toy"), or, in a separate group of subjects, to determine if the item was "old", "new", or "rotated". With this paradigm, the authors were able to show that when task demands were semantically driven, i.e., "toy" or "not toy", a

power function suggestive of pattern separation could be observed in the left hippocampus. When task demands emphasized spatial elements, i.e., "old", "new", or "rotated", the positive power function was found in the right hippocampus (Motley & Kirwan, 2012). Not only is this study important for showing that pattern separation can be observed in the human hippocampus, but it also emphasizes why mnemonic computation research is needed in humans. With the same stimuli, but with different instructions, different results were found: human intentions and goals can modulate what processes the brain performs.

High-resolution fMRI. The Montley and Kirwan (2012) paper used standard resolution fMRI in their experiment, but given the theory that pattern separation and pattern completion are performed in separate hippocampal subfields, hr-fMRI has been used in a number of studies. Despite its increased resolution, most experiments using hr-fMRI are still not able to anatomically differentiate the DG from CA_3; since these two areas are characterized by very different profiles of responding to changes in stimuli, the resulting BOLD signal could end up being a blend of DG- and CA_3-driven processes. The curve representing this combined signal would reflect a lower change in output for small stimulus changes than would be expected from the DG alone, resulting from the addition of pattern completion signals from CA_3. This curve would then shift to a larger change in output at higher levels of stimuli change where both DG and CA_3 pattern separate (Yassa & Stark, 2011). In a 2008 paper, Bakker, Kirwan, Miller, and Stark showed an overall pattern separation bias in a combined CA_3/DG region. In their study, like the previously described experiment, subjects were presented with either: a previously learned item, a new item, or a close lure to a previously learned item. In all

hippocampal subfields, they found repetition suppression (where repeats induce less lower activity than novels), but different responses to the lures across the subfields. Specifically, the CA_3DG showed greater activity for lures than repeats, with lure activity on par with that seen for novel items. In contrast, CA_1 showed the greatest activity for repeats, and lower activity for both lures and novel items. The authors concluded that CA_1 shows pattern completion and CA_3DG shows pattern separation (Bakker, Kirwan, Miller, & Stark, 2008).

Bakker et al.'s finding was replicated in a paper by Lacy, Yassa, Stark, Muftuler, and Stark (2011), which also showed the pattern where lure items show signal changes like novel items in CA_3DG, while showing signal changes more like repeats in CA_1. The Lacy et al. experiment additionally included two levels of similarity of lures (high and low similarity lures) in order to better evaluate the change functions of the subfields. In support of the current model, they found that in the CA_1 each step of change showed a corresponding linear change in BOLD activity, whereas in the CA_3DG ROI the first step of change (from repeat to high similarity) showed the largest difference in change of BOLD, and the next two steps up in change (to low similarity, and then to novel) showed little or no change from the BOLD activity to high similarity lures (Lacy et al., 2011).

The pattern separation response in the CA_3DG seems to be domain agnostic (Azab, Stark, & Stark, 2014; Huffman & Stark, 2014). CA_3DG responses to lure items, and original items presented in a new spatial location was similar to that of novel stimuli, and all were different from original items presented in their original location (Azab et al., 2014). All of the high-resolution studies on pattern separation to date use incidental encoding and retrieval tasks and provide strong evidence for the validity of the animal-

based models in humans. Further research is needed on how, or if, these patterns are affected by different task demands.

Not all studies have found differences in pattern separation and pattern completion between subfields. Kirwan & Stark (2007) conducted two experiments, one using faces, one using objects, to look at the effect of lures on the subfields. They found that the hippocampus was sensitive to hits, and both correct rejections of and false alarms to lures, but other regions did not show this sensitivity. However, they did not find a reliable pattern of activity between their two tasks, and found no differences in the activity of the subfields to the lure stimuli (Kirwan & Stark, 2007). Likewise, LaRocque et al. (2013) found evidence for pattern separation processes within the whole hippocampus, but not in the surrounding cortex, but found no difference in effects between subfields. In their study, they used RSA methods to quantify the degree to which the "neural signature" at encoding is reinstated at retrieval. More dissociable patterns between encoding and retrieval within the hippocampus predicted greater memory success, while the opposite was true in MTL cortex, where more similarity of patterns was a greater predictor of memory (LaRocque et al., 2013).

Sequence Processing

There are some who suggest that the spatial processing that seems to be particular to the hippocampus (for reviews of spatial processing in the hippocampus see: Eichenbaum, 2000; Moser, Kropff, & Moser, 2008) might be an extension of one of the hippocampus's more basic functions: the representation of ordered sequences (Eichenbaum, 2013). Under this theory, the hippocampus serves to encode and hold information in order, both spatially and temporally. While spatial information makes up a

large part of the context of memory, the events that happen before and after are equally important to memory context, and much of the research on sequence encoding/retrieval focuses on the temporal elements of memory. In fact, while spatial memory is now the emphasis for much of the work in rodent episodic (if it can so be called) memory, the first type of learning that was determined to be hippocampal-dependent was trace conditioning. In this type of conditioning the unconditioned and conditioned stimuli, unlike most other forms of conditioning, are presented with a time gap between them (Bangasser, 2006). While not direct evidence for the hippocampus's role in learning order, it shows clearly that the hippocampus plays an essential role in binding items across time, as well as space. Furthermore, older adults, who often show marked deficits in hippocampal functioning (Tsukiura et al., 2010; Yassa et al., 2010; Yassa, Mattfeld, Stark, & Stark, 2011), have showed impairments in sequence encoding (Allen, Morris, Stark, Fortin, & Stark, 2015), a further implication of the likely role that the hippocampus plays in sequence learning.

Human research on the representation of sequences in the hippocampus. ***Standard resolution fMRI.*** Only a few fMRI studies of human sequence representation in the hippocampus have been conducted and have resulted in conflicting results. Two studies have shown that hippocampus engagement predicts subsequent recall of information that included a temporal component. Staresina & Davachi (2009) showed that recall for item-color pairs that were presented discontiguous in both time and space showed the greatest reliance on the hippocampus, where correctly recalled items elicited much greater activity in the hippocampus as compared to items that were forgotten. Pairings that were

discontiguous only spatially also relied on hippocampal activity at encoding, but the difference was not as great as the spatiotemporally discontiguous items (Staresina & Davachi, 2009). Tubridy and Davachi (2011) showed similar results; in their experiment subjects were shown sequences of words, with each presentation temporally spaced from each other. Those words that later went on to be recalled with the correct temporal order showed greater hippocampal activity during encoding than forgotten words and words whose orders were forgotten (Tubridy & Davachi, 2011).

In opposition to the above studies, another study failed to find differential hippocampal engagement for the recall of order. St. Jacques, Rubin, LaBar, and Cabeza (2013) used photos from subjects' own experiences to test their recall for the order in which they had initially encountered the items/locations shown in the photos. These researchers expected areas that were sensitive to temporal order would show greater activation when the task required subjects to recall the sequential order of items encountered closer in time than when it required them to recall the sequential order of more distantly spaced items. While they found PHC differences based on the duration of the gaps between items, they did not find any differences in the hippocampus (St. Jacques, Rubin, LaBar, & Cabeza, 2008). A possible explanation for this finding is that the activity in the hippocampus is equally engaged for recall of any temporal information, but might still exhibit large differences between its patterns of activity depending on the temporal difference between experiences. An MVPA or representational similarity analysis approach might be employed to test this hypothesis.

Less direct, but still strong evidence for the hippocampal encoding of sequence information comes from work regarding mismatch detection. This topic will be discussed

in much greater length below, but mismatch detection refers to an increase in activity exhibited when a stimulus is encountered that mismatches with what would have be expected based on one's past experiences. When subjects are presented with an out-of-order second half of a previously learned sequence, a strong mismatch signal is seen in the hippocampus, even when compared to signals measured in response to an entirely new sequence with all novel items (Kumaran & Maguire, 2007a). This mismatch signal is elicited even when a sequence has been encountered only once, indicating that the sequence is encoded and later predicted as a matter of course in the hippocampus.

High-resolution fMRI. Only one study to date has looked at human hippocampal subfields' contributions to sequence functions. Azab, Stark, and Stark (2013) sought to uncover any differences in pattern separation activity between spatial and temporal sequences. Using similar methods to their other work with pattern separation, subjects were presented with lure conditions to previously learned spatiotemporal patterns: ordered sequences of items appearing in specific locations on a checkerboard. The lures either switched the temporal order (items appeared in the same place, but in a different order than the learned pattern), switched the spatial location of the items (items appeared in the same order but in different locations), or switched both (items appeared in a different order and in different locations). All lure conditions showed the same amount of pattern separation signal in the combined CA_3/DG ROI as well as in the CA_1 ROI, though this region showed a trend towards additional decrease in signal for the combined lure condition (Azab et al., 2014). This paper provides further support for the

re-association of space and sequence, as the results indicate that these two manipulations are treated as part of the same process.

Mismatch Detection

Theoretical model. An offshoot of the research into hippocampal pattern separation/completion, prediction, and sequence encoding is the theory of mismatch detection. As described earlier, mismatch detection refers to a signal that indicates when a stimulus mismatches what would have been expected based on previous memory. Theoretically, mismatch detection would be an important part of a network designed to make predictions about the future; in order for this system to continue to make predictions with high validity, it is necessary to continually monitor for any errors or changes and update one's memories, and hence future predictions, accordingly. Likewise, in the context of pattern separation and completion, if the results of pattern completion do not represent the correct current state of the external world, a signal is needed to indicate that this is a novel event to encode/pattern separate.

Evidence for mismatch detection in the hippocampus. ***Standard resolution fMRI and direct recording.*** As was mentioned in the sequence section, this mismatch signal has been observed when an item is presented in the incorrect order relative to a previous presentation. In a study using standard resolution fMRI, the signal from the hippocampus was maximally elevated when the order of a sequence was violated, above even the activation level obtained in response to the presentation of a new sequence (Kumaran & Maguire, 2007a). This associative novelty, rather than absolute novelty, seems to be unique to the hippocampus. Direct recordings from epileptic

patients demonstrated an increase in low power oscillations in the hippocampus for associative novelty, reinforcing the fMRI findings (Chen et al., 2013).

High-resolution fMRI and biological model. Mismatch detection is thought to depend on a comparator mechanism within the hippocampus, with the CA_1 as the most likely region (Kumaran & Maguire, 2007b), although the CA_3 might also perform some comparator functions (Vinogradova, 2001). The CA_1 receives inputs both from the CA_3 and its pattern completion outputs, as well as directly from layer 3 of the ERC, which can convey updated information from sensory reality. In this position, the CA_1 can compare the outputs of pattern completion, which can be thought of as a memory-based prediction, with emerging reality (Kumaran & Maguire, 2007b). If the CA_1 detects a mismatch, it can signal via the feedback loop with the NAcc and VTA to trigger new memory encoding (Lisman & Grace, 2005). In support of this model hr-fMRI studies have found that the CA_1 does show increased activity when presented with an item that mismatches with what was previously paired with a cueing stimulus (Chen, Olsen, Preston, Glover, & Wagner, 2011; Duncan, Ketz, Inati, & Davachi, 2012).

The magnitude of the mismatch response within the CA_1 is modulated by prediction strength (Chen, Cook, & Wagner, 2015). Subjects were shown sequences of four items, either once or four times. On the next and final presentation the third and fourth items were swapped. In the CA_1 this order violation resulted in greater activity for those sequences seen multiple times than for those seen only once before the violation. The authors concluded that the increased prediction strength expected during the viewing of a sequence seen repeatedly caused the increase in CA_1 activity to the violations. Furthermore, in a separate behavioral experiment the authors showed that

items that violated their predicted presentation order resulted in increased recognition during a subsequent memory test. This result is the first to show that mnemonic mismatches can lead to improved memory, as predicted by the Hippocampal-VTA Loop model (Lisman & Grace, 2005). Despite providing the first evidence for the behavioral effect of mismatch detection, the difference between hit rates between repeat and violation conditions was fairly modest. Further work replicating this effect is called for.

Match/mismatch detection modulated by motivation. ***Standard resolution.*** While it is often thought that the mismatch signal observed in the hippocampus reflects the initial detection of a mismatch by the hippocampus, the observed signal increase might, rather, be due to the dopaminergic inputs from the VTA (Lisman & Grace, 2005). Evidence for this theory comes from a study by Murty and Adcock (2014) that showed that reward-motivated, task-irrelevant, associative mismatches elicited greater hippocampal signals, and this increase was predicted by increased connectivity of the VTA reward circuitry (Murty & Adcock, 2014). However, studies using direct neural recordings have shown that changes in oscillatory power can be detected as soon as a mismatching stimulus is presented (Chen et al., 2013). Thus, while the strong effects seen in BOLD data might at least partially be influenced by dopamine, the initial mismatch response still likely originates in the hippocampus.

In line with the hippocampal mismatch signal being modulated by reward states, a number of studies, both with high and standard resolution fMRI, have shown shifts to match enhancement when subjects are engaged in a task pertaining to the detection in changes in stimuli. Duncan, Curtis, and Davachi (2009) showed match enhancement in the whole hippocampus when subjects were asked to determine if the items had

changed after a delay (Duncan, Curtis, & Davachi, 2009). Match enhancement was also seen in more complex tasks, such as a study by Hannula and Ranganath (2008), which required subjects to mentally rotate a scene and make decisions regarding changes in items' locations. Match enhancement was greatest when they were presented with a scene that matched the rotated view, and then progressively less as either and/or both the position or items of the scenes were mismatching (Hannula & Ranganath, 2008).

High-resolution. Dudukovic, Preston, Archie, Glover, and Wagner (2011) showed in an hr-fMRI study that all subfields exhibited match enhancement in a task that required subjects to explicitly determine if there was a mismatch (Dudukovic, Preston, Archie, Glover, & Wagner, 2011). Duncan, Ketz, Inati, and Davachi (2012) expanded on their earlier findings with an hr-fMRI study that had both an implicit and explicit task. They showed that during the implicit task, the CA_1 showed mismatch detection that was graded by the number of changes present in a scene, more change elicited more signal increase. During the explicit task, the opposite pattern was found: graded match enhancement, with less signal change for the most stimuli changes (Duncan et al., 2012). The graded change in the CA_1 signal in response to change in stimuli is in line with the pattern separation model, which places CA_1 as the diagonal line of the change graph, neither performing pattern separation or completion, rather tracking the total amount of change.

CHAPTER 2

Dissociable Profiles of Generalization/Discrimination in the Human Hippocampus during a Visual Associative Memory Retrieval Task

One fruitful approach for characterizing mnemonic processing within distinct hippocampal subfields has been to measure neural responses to stimuli that closely, but not exactly, match a previously encountered item or association. If a region's response to these 'close lure' stimuli mimics its response to novel stimuli, the region is said to show a discrimination effect. This is sometimes referred to as pattern separation (Deuker, Doeller, Fell, & Axmacher, 2014; Yassa & Stark, 2011), since the close lure must be encoded in such a manner as to minimize its representational overlap with the previously learned stimulus, thus preventing catastrophic mnemonic interference (Kesner, 2013a, 2013c). Previous high-resolution functional MRI (hr-fMRI) studies have predominantly observed mnemonic discrimination effects in the CA3 and dentate gyrus subfields of the hippocampus (Bakker, Kirwan, Miller, & Stark, 2008; Lacy et al., 2011). Because most hr-fMRI protocols lack the resolution to differentiate between these regions (Carr et al., 2010), we will follow the convention of referring to them as a combined CA3DG region.

Mnemonic generalization effects, on the other hand, have been reported in the downstream hippocampal subfield of CA_1, which responds similarly to close lure stimuli and exact repeats of a past stimulus (Bakker et al., 2008; Lacy et al., 2011). Such generalization possibly reflects pattern completion mechanisms (recall based on degraded stimuli (Kesner, 2013b, 2013c)). Alternatively, the similarity in responses to close lures could be the result of linear changes in neural activity in response to

changing stimuli (Lacy et al., 2011; J. K. Leutgeb et al., 2005), that does not result in significant differences for small variations. However, extant fMRI-based demonstrations of this dissociation between CA_3DG and CA_1 rely on incidental recognition paradigms (i.e., tasks requiring no actual memory judgments) making it challenging to evaluate the degree to which mnemonic encoding and/or retrieval operations drive these respective response profiles (Deuker et al., 2014; Liang & Preston, 2015).

Several fMRI studies have examined hippocampal sensitivity to whether events match or mismatch one's memory-based expectation (e.g., associative novelty)(Kumaran & Maguire, 2006, 2007a; for reviews: Kumaran & Maguire, 2007b, 2009) and reported effects in the CA_1 region (Chen et al., 2015, 2011; Duncan et al., 2012). Since CA_1 receives direct inputs from CA_3, putatively conveying associative information derived from CA_3-mediated pattern completion, and from entorhinal cortex, putatively conveying information about sensory reality (Amaral, 1993; Amaral & Witter, 1989), CA_1 is uniquely situated to compare these two signals and evaluate whether one's memory-based expectation is consistent with the current state of the world (Vinogradova, 2001). Furthermore, CA_1 connects to a larger reward/motivation system that allows for memory modulation in response to both external cues (such as mnemonic mismatches) and internal drives (Lisman & Grace, 2005). Indeed, when subjects are explicitly tasked with judging whether a probe stimulus matches a past memory, match enhancement (repeat > novel) signals are typically observed throughout the hippocampus (Duncan et al., 2009; Hannula & Ranganath, 2008) and are especially pronounced in CA_1 (Dudukovic et al., 2011; Duncan et al., 2012). Some explicit tasks, however, find mismatch enhancement (Chen et al., 2011; Duncan et al., 2012).

We conducted an hr-fMRI study to determine if the previously described profile of mnemonic discrimination in CA_3DG and generalization in CA_1 persists in an explicit associative memory recall task and, furthermore, how the demands of the task affect these patterns.

Hypotheses

- The hippocampus, and the CA_1 particularly, will exhibit match enhancement.
- CA_3DG will only show enhancement for repeated stimuli, and not lure stimuli, while CA_1 will not discriminate lures from repeats as robustly.
- CA_1, as a part of the motivation network, will show sensitivity to task demands while CA_3DG will not.

Methods

Participants

Twenty-nine right-handed healthy members of the University of California, Los Angeles (UCLA) community participated in this study (15 females; mean age = 20.79 years, range 18 to 28 years). Participants consented and received compensation in compliance with the UCLA Institutional Review Board. Four subjects' data were excluded from the dataset due to issues affecting data quality. In total, 25 participants make up this data set (13 females; mean age = 20.6; range from 18 to 24 years).

Task Design

Our task was comprised of two parts: a learning phase and a testing phase (Figure 2.1). During the learning phase, subjects were presented with 168 randomly generated name/image associations, each consisting of a celebrity's name and a visual image of an object, an animal, or a building/landmark. Each pair was displayed for 6

sec, and subjects were told to explicitly form an associative memory connecting the name with the image. The session included two study rounds, such that each pair was encountered twice. On the following day, participants returned for an fMRI scanning session, during which their memory for the associations was tested. Each trial began with the presentation of a celebrity name, and subjects were to mentally recall the associated image and rate the strength of their memory on a 4-level scale: “Strong” - “No” memory). After an 8 sec delay, a probe image was presented for 2 sec. Probe stimuli could either be the previously learned image associate (*targets*), a similar image to the learned associate (i.e., a different exemplar of the same item; *foils*), or a very different image than the learned associate (i.e., an exemplar from a different visual category). For this final class of probes, half had been previously encountered during the learning phase, but associated with a different celebrity name, withheld from the testing (*mispaired*) and half were presented for the first time during the test phase (*novels*). Subjects were to make a button-press response indicating whether they thought the probe was an: “Exact match” (the correct response for targets), “Similar” (correct for lures), “Very Different” (correct for novels and mispaired probes), or “Do Not Know” (a response option included to discourage guessing). Subjects performed a total of 144 memory trials (48 targets, 48 lures, 24 mispaired, and 24 novels; one presentation of each 168 name cue, less the 24 pairing withheld for use as the mispaired trials). In an effort to prevent mind-wandering during the inter-trial interval, which can be associated with hippocampal activity (Stark & Squire, 2001), subjects performed an active baseline task for 8 sec between trials, requiring them to report the right/left direction of a series of arrow stimuli.

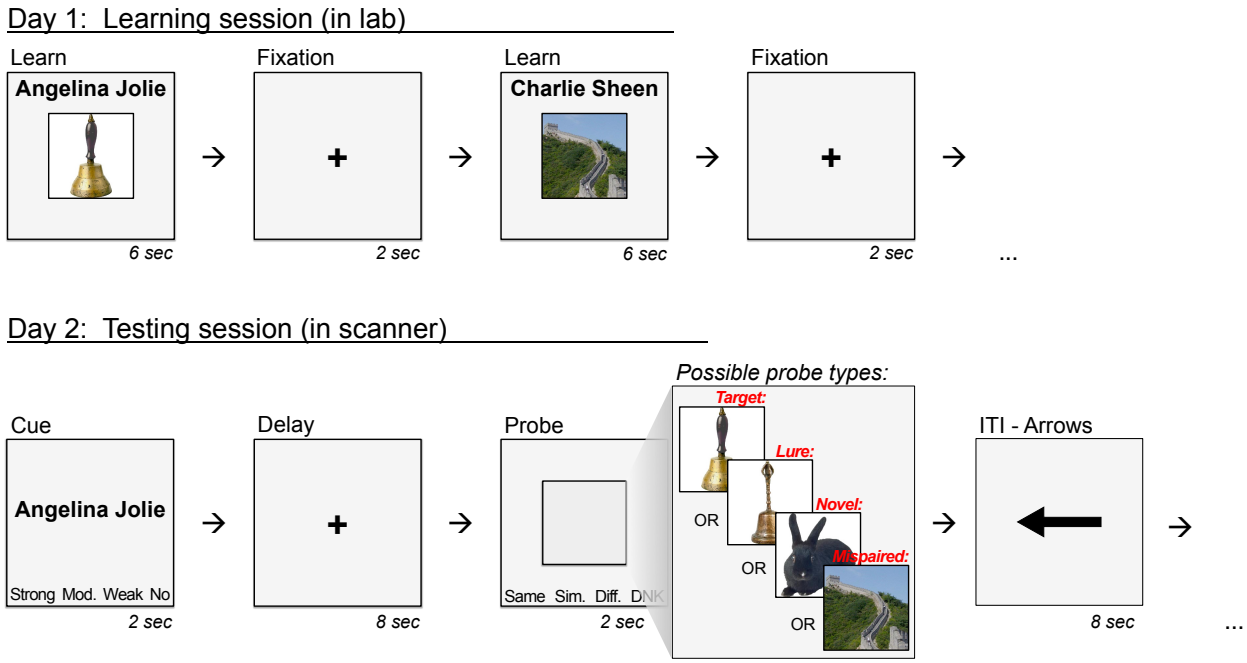


Figure 2.1: Experimental Design.

MRI Data Acquisition

Neuroimaging data were acquired on a 3.0 T Siemens Tim Trio MRI system at the UCLA Staglin IMHRO Center for Cognitive Neuroscience. Functional scans used a T2*-weighted gradient-echo planar imaging (EPI) pulse sequence (repetition time (TR) = 2500 ms; echo time (TE) = 28 ms; GRAPPA acceleration factor = 2; 38 slices aligned off angle to the hippocampal long axis; 1.7 x 1.7 x 1.8 mm voxel size; 8 runs; 146 volumes per run). A T1-weighted magnetization-prepared rapid-acquisition gradient echo (MPRAGE; 240 sagittal slices, 1 x 1 x 1 mm voxels) acquired anatomical images of the whole brain. A T2*-weighted high-resolution anatomical scan of the hippocampus (28 slices aligned perpendicular to the long axis of the hippocampus, 0.4 x .04 x 1.8 mm voxels) allowed for subfield ROI creation. MRI-compatible LCD goggles presented visual stimuli and a four-button box recorded subjects' responses.

ROI Creation, Registration and Pre-processing

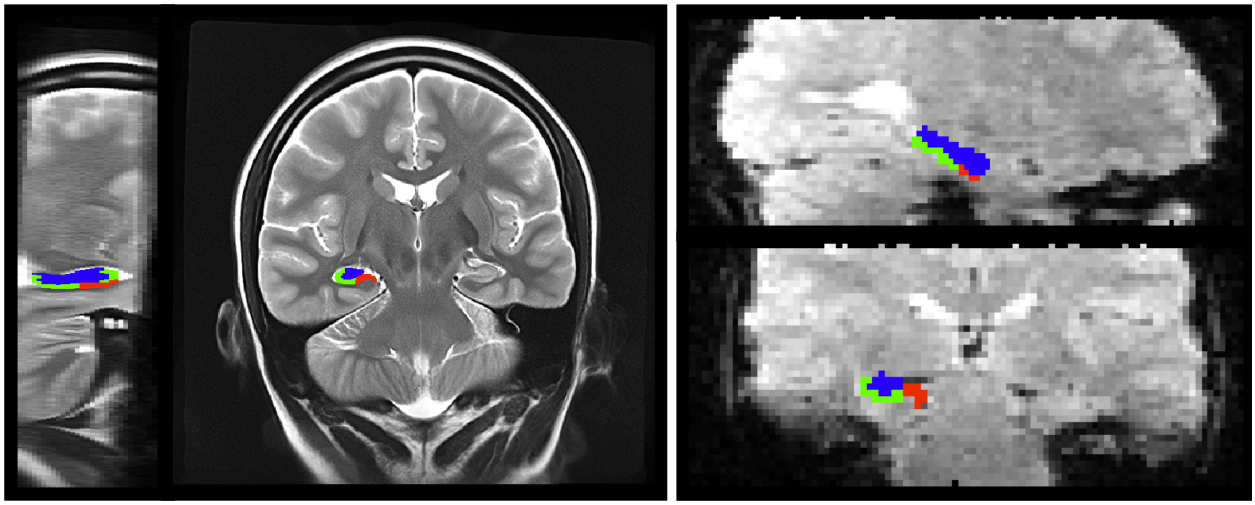


Figure 2.2: (A) Example hand-drawn hippocampal subfield ROIs (only left hemisphere shown) overlaid on the template high-resolution coronal structural image. (B) Example subject's ROIs (only left hemisphere shown) following transformation to native EPI space, overlaid on a single EPI volume.

Hippocampal ROIs were hand drawn for each subject on their high-res coronal structural image following conventional region boundaries (Duvernoy et al., 2013; Frankó et al., 2014), Figure 2.2A. All fMRI BOLD data extraction analyses were performed in native EPI space, and several registration steps using Advanced Normalization Tools (ANTs) (Avants et al., 2010, 2011) moved the ROIs into this space, Figure 2.2B. Template EPI space created for each subject based on exemplar images from each of the eight scanner runs accommodated any small drifts resulting from subject motion over the course of the scanning session, and transforms of each EPI run to the template were calculated. Both the template and the high-res hippocampal image were aligned to subject's MPRAGE image. ROIs created in the high-res anatomical space were moved directly to EPI space by applying the appropriate set of transformation parameters. An additional transform of subject MPRAGE to a study specific whole-brain template moved EPI images to a common space for the creation of group maps.

The data were slice time corrected, and motion and magnetic field distortion corrections were performed concurrently using FSL's (Jenkinson, Beckmann, Behrens, Woolrich, & Smith, 2012) FLIRT (Jenkinson, Bannister, Brady, & Smith, 2002) (FMRIB's motion correction linear image registration tool) and field unwarping tool (FUGUE). To prevent the blending of activity across neighboring hippocampal subfields, the data used in the ROI analyses were not spatially smoothed. However, when generating group-level activity maps, a 5mm Gaussian FWHM smoothing kernel was applied the EPI data to accommodate across-subject variability in subfield boundaries.

fMRI Data Analysis

All statistical modeling was performed with FSL's FEAT (FMRI Expert Analysis Tool). Our model investigating activity during probe presentation consisted of six regressors, represented by 2-sec boxcars convolved with a gamma hemodynamic function, one each for successfully classified probes based on probe type (targets, lures, mispairs, novels), one for all incorrect trials collapsed across trial types, and one for the cue period. The ITI period remained unmodeled and served as an "active baseline" (Stark & Squire, 2001). Six standard motion parameters, as well as their first and second derivatives, were included as regressors of no interest. This resulted in activity parameter estimates (betas) for each subject for all task conditions occurring during each run (since subjects' responses drove accuracy, the number of trials per condition varied, and not all conditions occurred in every run).

For subfield ROI analyses, betas were converted to percent signal change (Mumford, 2007), representing each condition's change from baseline. Signal change estimates within each ROI were extracted for conditions in each run, and a weighted

average based on the number of trials per run was then calculated to generate across-run mean signal change estimates. For whole hippocampus analyses, group masks were created using FSL's multilevel FEAT tools: first-level analyses were performed on smoothed data; then fixed effects were calculated to combine parameter estimates across runs; finally, mixed effects group contrasts were calculated across subjects using FSL's FLAME1 method. Clusters were thresholded at $p<0.05$ corrected, using a whole hippocampus mask for small volume correction.

Results

Behavioral Results

During the cue period when subjects responded and recalled the paired image (*mean* = 83.27% of trials, *SD* = 18.01%) they used the three levels of memory strength roughly equally ($F_{(2,48)}$ = 2.211, *MSE* = 844.651, *p* = 0.121). The "do not know/remember" rating (*mean* = 16.72% *SD* = 18.01%) was used less than the other three ratings on average (*mean* = 26.87%, *SD* = 6.21%; $t_{(24)}$ = 2.35, *p* = 0.027). Memory rating significantly predicted success at the probe, with a strong linear relationship between a larger percent of correct trials at higher memory strengths ($F_{(1,24)}$ = 10.274, *MSE* = 504.575, *p* = 0.004). The increase in accuracy by memory strength is also seen in reaction time (RT). On trials that subjects gave a

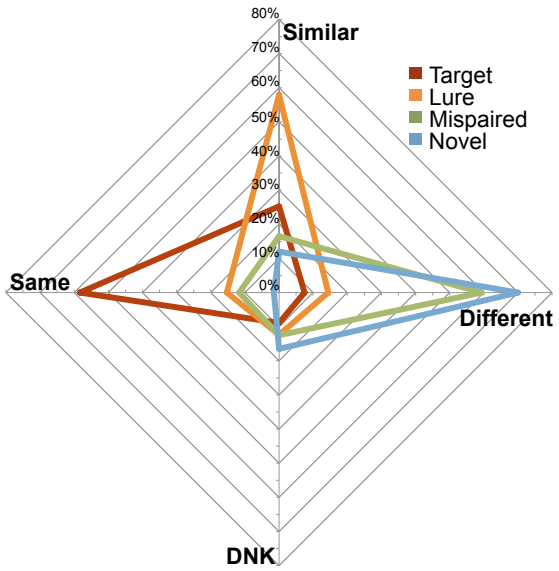


Figure 2.3: Behavioral responses breakdown by condition types.

"high" memory rating to the cue, RT to the probe (*mean* = 1.35, *SD* = 0.18) was significantly faster than to the probe for trials given the other three memory ratings on average (*mean* = 1.56 sec, *SD* = 0.20; $t_{(24)} = 4.73$, $p < 0.001$), which did not differ from each other ($F_{(2,44)} = 0.071$, *MSE* = 0.025, $p = 0.932$). Subjects were much quicker when reporting a high memory rating for the cue's pair (*mean* = 1.351 sec, *SD* = 0.18 sec) than the other three on average (*mean* = 1.56 sec, *SD* = 0.020 sec), which did not differ from each other ($F_{(1,22)} = 0.012$, *MSE* = 0.039, $p = 0.914$).

Subjects correctly determined probe type on 58.3% (*SD* = 16.5%) of trials, and while accuracy differed between conditions ($F_{(3,72)} = 3.748$, *MSE* = 0.022, $p = 0.015$), only identification of novel images was greater than the other conditions (novel: *mean* = 70.00%, *SD* = 19.17%; target: *mean* = 58.33%, *SD* = 21.37%; $t_{(24)} = 2.62$, $p = 0.015$; lure: *mean* = 57.92%, *SD* = 18.34%; $t_{(24)} = 2.68$, $p = 0.013$; mispaired: *mean* = 59.33%, *SD* = 20.42%; $t_{(24)} = 3.28$, $p = 0.003$); there was no difference between the remaining conditions ($F_{(2,48)} = 0.057$, *MSE* = 0.023, $p = 0.945$) (Figure 1B). Furthermore, on correct trials, subjects' RT varied by probe type ($F_{(3,72)} = 25.659$, *MSE* = 0.009, $p < 0.001$), reflecting faster RTs for detecting target probes (*mean* = 1.24 sec, *SD* = 0.17) or novels (*mean* = 1.29, *SD* = 0.19) than either lures (*mean* = 1.40 sec, *SD* = 0.16; vs. targets: $t_{(24)} = 6.87$, $p < 0.001$; vs. novels: $t_{(24)} = 4.81$, $p < 0.001$) or mispaired probes (*mean* = 1.45 sec, *SD* = 0.24; vs. targets: $t_{(24)} = 6.59$, $p < 0.001$; vs. novels: $t_{(24)} = 5.97$, $p < 0.001$). See Figure 2.3 for a breakdown of responses by condition.

Hippocampal Activity Tracks Memory Strength

Many regions showed a linear increase in BOLD activity reflecting subjects reported memory strength at the cue. Significant clusters reflecting strong > moderate >

low > no memory were found bilaterally in the precuneus and cingulate cortices and extended throughout visual association regions. Clusters were also found in left prefrontal cortex (both medial and lateral) and bilateral caudate, insula, middle temporal gyrus, and hippocampus. Figure 2.4A shows significant clusters for this effect within the hippocampus. The relatively large set of regions exhibiting modulation by memory strength may reflect a number of processes involved in making this memory evaluation, including memory retrieval in the hippocampus, neocortical reinstatement in visual associative regions (as subjects were instructed to vividly maintain the paired associate in mind during the delay period), and strength evaluation in the medial prefrontal cortex.

Left Hippocampus Shows Associative Match Enhancement

In order to confirm previous findings that the hippocampus shows differential responses to novel items relative to, in this case, correctly paired target items, we bi-directionally contrasted these two probe conditions for correct trials. Consistent with findings of stronger activation for items that match expectations in a task (Dudukovic et al., 2011; Duncan et al., 2012), we found a significant cluster within the left hippocampus (58 continuous voxels, Z-max = 3.44 at [-24 -22 -18], p = 0.007, small volume corrected) that showed greater activity for target probes than for novel probes (Figure 2.4B). No clusters emerged for novels > targets, indicating that hippocampal novelty detection and/or repetition suppression effects observed in prior studies using incidental recognition tasks (e.g., Bakker et al., 2008; Lacy et al., 2011) may not be characteristic of explicit retrieval tasks like ours. Furthermore, our mapwise analyses revealed no significant findings in the right hippocampus. Thus we elected to focus our subfield ROI analyses exclusively on the left hippocampus, which may have shown

more robust effects in our task due to subjects' reliance on verbal/semantic processing to promote the memorability of name-image associations. Both the left CA_1 ($F_{(1,24)}$ = 7.963, *MSE* = 0.020, p = 0.009) and CA_3DG ($F_{(1,24)}$ = 13.366, *MSE* = 0.020, p = 0.001) showed significantly greater activity to correct targets than novel probes.

Dissociable Profiles of Discrimination and Expectation Dependency

Within the context of match enhancement, we next wanted to examine the degree to which these two subfields exhibit activity patterns reflecting generalization or discrimination, and the expectation dependency of these patterns. The relative activity of the lure probes compared to target and novel probes allows us to determine if a region generalizes (i.e., shows comparable activity for lure and target images) or

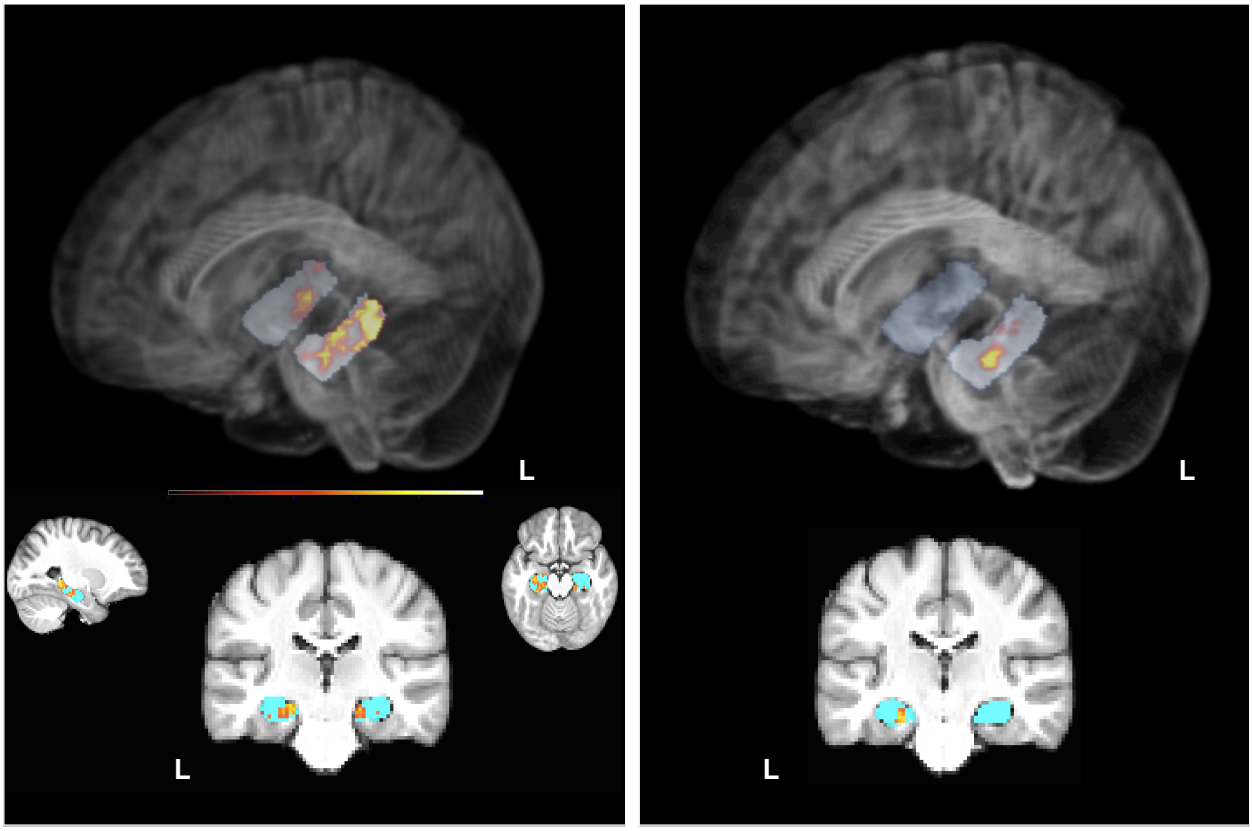


Figure 2.4: (A) Group-level linear effects for memory strength to cue stimuli within a whole hippocampus mask. (B) Group-level contrast of target and novel probe.

discriminates (i.e., shows comparable activity for lure and novel images). The mispaired, but pre-exposed, probe images can provide insight into the expectation dependency of a region's match enhancement effect. A region could be expectation-dependent, showing increased activity only for correctly paired images (i.e., comparable activity for mispaired and novel images), or expectation-independent, where any previously seen image elicits match enhancement (i.e., comparable activity for mispaired and target images). Figure 2.5 summarizes schematically four hypothetical profiles of dependency and discrimination.

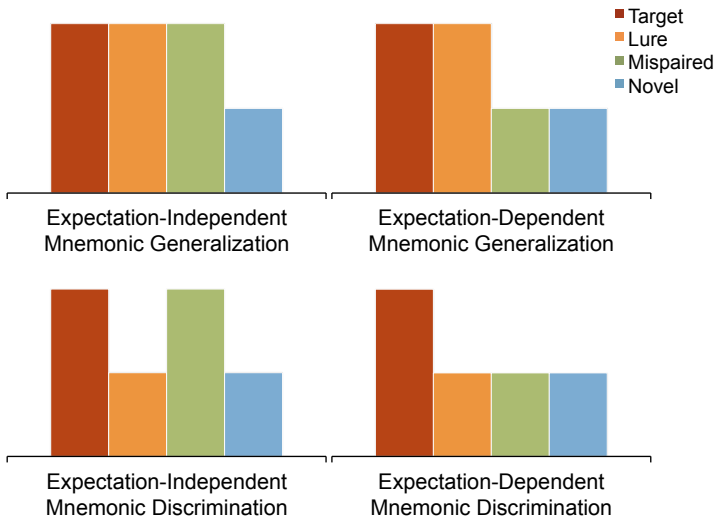


Figure 2.5: Four hypothetical response patterns for combinations of expectation dependency and mnemonic discrimination

In order to increase the sensitivity of our ROI analyses, we first identified a subset of task-responsive voxels within each region, based on a statistically independent analysis of the incorrect trials (which were not included in our analyses of interest). For each subject, we identified the top 25% of voxels within each ROI that showed the greatest response to probes on incorrect trials (collapsed across probe conditions). Probe activity estimates for correct trials were then extracted from the resulting ROIs for all probe conditions (Figure 2.6). To evaluate whether left CA_3DG and CA_1 showed dissociable profiles of responding, we first tested for a Region x Probe

Condition interaction, which showed a marginally significant effect ($F_{(3,72)}$ = 2.643, *MSE* = 0.049, p = 0.056). Given that our previous analyses showed greater activity for target probes in both subfields, we next sought to evaluate whether the two subfields would show dissociable effects when only considering their responses across the other three probe conditions, which each involve some degree of mnemonic mismatch to the previously learned associate. This interaction proved significant ($F_{(2,48)}$ = 3.452, *MSE* = 0.056, p = 0.040), indicating that CA_3DG and CA_1 show a differential responsivity to lure, mispaired, and novel images.

Differential Responses to Lure Probes

In both regions, we next compared responses to lure probes to those elicited by target or novel probes. In CA_1 lures did not differ from targets ($F_{(1,24)}$ = 0.139, *MSE* = 0.032, p = 0.713), while they did differ from novels ($F_{(1,24)}$ = 5.734, *MSE* = 0.050, p =

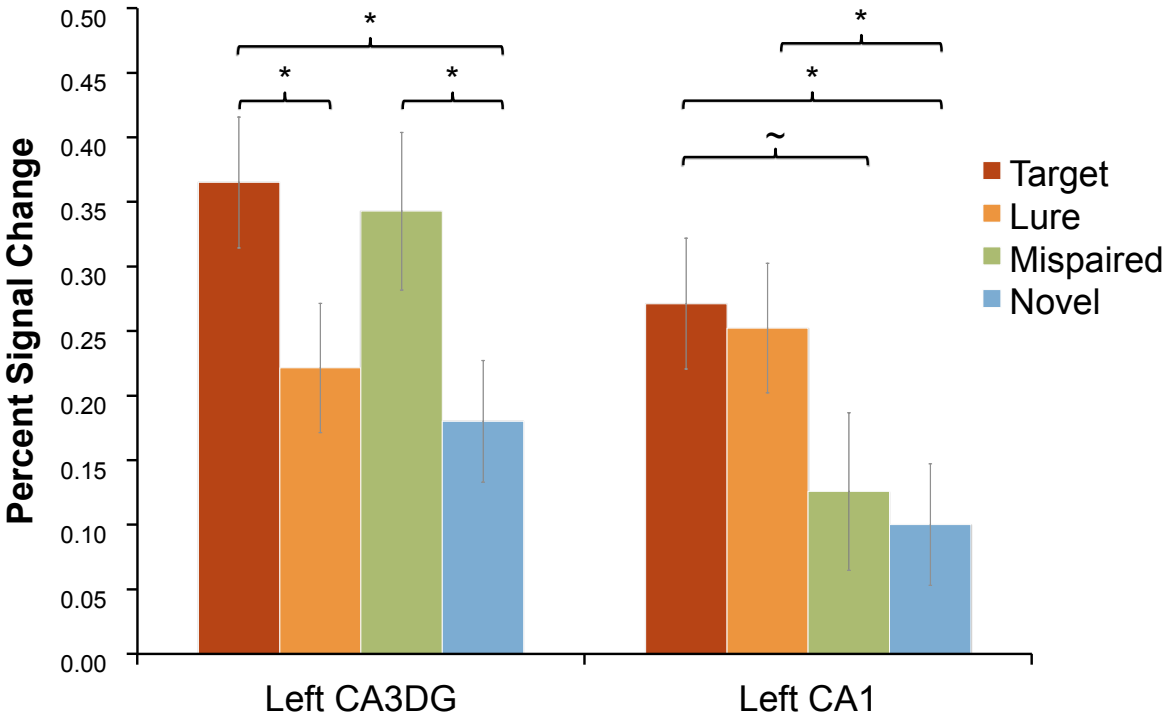


Figure 2.6: Mean percent signal change within left CA_1 and CA_3DG ROIs as a function of probe type

0.025). In contrast, in CA_3DG lures differed significantly from targets ($F_{(1,24)}$ = 6.063, *MSE* = 0.043, p = 0.021), but showed no difference from novels ($F_{(1,24)}$ = 0.440, *MSE* = 0.049, p = 0.513), Figure 3. This finding comports with results from prior studies that found sensitivity to subtle visual changes in CA_3DG and sensitivity to larger changes in CA_1 (Bakker et al., 2008; Lacy et al., 2011).

Differential Responses to Mispaired Probes

We next contrasted mispaired probes with target and novel probes. CA_3DG showed increased activity for mispaired images compared to novel ($F_{(1,24)}$ = 11.935, *MSE* = 0.028, p = 0.002), but did not differ from targets ($F_{(1,24)}$ = 0.134, *MSE* = 0.046, p = 0.718), reflecting a generalized preference for familiar stimuli, regardless of whether the specific probe image was expected (i.e., targets) or unexpected (i.e., mispaired) in the context of a given trial. CA_1, however, showed no difference between mispaired and novel probes ($F_{(1,24)}$ = 0.127, *MSE* = 0.065, p = 0.725), but showed a trend towards mispaired probes evoking less activity than targets ($F_{(1,24)}$ = 3.829, *MSE* = 0.069, p = 0.062). Thus, CA_1 showed high activity on trials where the probe stimulus was either an exact match (targets) or near match (lures) to one's memory-based expectation, but lower activity on trials where the probe's content was unanticipated. Importantly, this effect was observed regardless of whether the unexpected probe stimulus was truly novel or whether it was merely contextually novel (previously seen in the context of a different name-image pairing).

Discussion

The findings here conform with prior fMRI work showing that the CA_3DG subfield of human hippocampus responds to lure images as it would to novel images

(discrimination), whereas CA_1 responds to lures as it would to target images (generalization) (Bakker et al., 2008; Lacy et al., 2011). However, unlike previous studies which observed these dissociable subfield response profiles in the broader context of repetition suppression effects (decreased activity to targets), here we document a similar dissociation in the context of match enhancement effects (increased activity to targets). Furthermore, the activity increase to targets we observe in both subfields aligns with earlier work showing increased activity to matching conditions when subjects were explicitly tasked with evaluating the whether the probe stimulus matched or mismatched one's expectations (Dudukovic et al., 2011; Duncan et al., 2009, 2012; Hannula & Ranganath, 2008). These studies found such match enhancement exclusively in CA_1; the additional CA_3DG results here may potentially reflect additional "recall-to-reject" (Rotello & Heit, 2000) processes required to discriminate matches (targets) from incorrect pairings (mispaired), as well as from novel and lure stimuli.

Importantly, our results demonstrate dissociable hippocampal response profiles as a function of condition. CA_3DG discriminates old visual stimuli from new stimuli, irrespective of the task-relevance expectations. That is, it shows the same degree of match enhancement to all previously encountered images, regardless of their associative relationship to the name recall cue. In contrast, CA_1 responds to incorrectly paired targets as it would to novel probes, while it generalizes to different exemplars of the expected image. This dissociation, along with the switch to match enhancement seen in most other studies that have used explicit memory judgment tasks (although see Chen et al., 2011 for one notable exception), seems to reflect top-down, goal-

directed, influences on CA_1 responses. The seeming importance of task demands to CA_1 responses would indicate that the fMRI signal in CA_1 in response to mismatches seen in other studies likely reflects the end result of neuromodulatory inputs from the downward arc of the hippocampal-VTA loop, which conveys motivation and goal information in addition to the hippocampal mismatch signal, (Lisman & Grace, 2005), rather than an actual mismatch response. However, as our experiment lacks a manipulation of task demands, we cannot definitively conclude that this reflects a truly task-specific effect.

Our results further emphasize the challenges inherent in attributing the differential lure sensitivity of the CA_3DG and CA_1 subfields to pattern separation and pattern completion mechanisms, given that mechanisms are typically associated with encoding and recall processes, respectively (Hunsaker & Kesner, 2013). The active judgments about images' relationship to cues presented in our task likely require the employment of a "recall-to-reject" strategy. Within the context of this strategy, increased activity within the CA_1 to close lures could actually reflect pattern completion. However, the increased response to targets and mispaired images by the CA_3DG region does not fit with a pattern separation model, as this discrimination response likely does not reflect selective encoding of these already-familiar images. On their own, the CA_3DG responses seem to indicate selective reactivation to previously encountered stimuli.

Without more conditions presenting additional degrees of perceptual dissimilarity for lure images, we cannot say with certainty that the increased activity to both targets and lures observed in CA_1 reflects insensitivity to lures. As previous research shows that CA_1 linearly tracks the amount of change in stimuli and environments (Lacy et al.,

2011; J. K. Leutgeb et al., 2005), we possibly would have observed a more graded effect had we added another level of lure dissimilarity (e.g., objects that share the same general category as the target image). Finally, our experiment lacked power to allow an examination of the potentially informative responses of the hippocampal subfields to lure trials that subjects either misidentified as “same” or “different”, nor did we have sufficient trials to examine the effect of subjectively reported memory strength on our pattern of results. Future efforts more focused on these questions could illuminate hippocampal subfield contributions to subjective experiences leading to over-generalization and/or forgetting, as well as the modulation of mismatch responses by prediction strength (Chen et al., 2015).

Conclusion

Here we show diverging sensitively to changes in stimuli and task demands in the responses within the left CA_1 and CA_3DG subfields of the human hippocampus. Both regions showed greater responses to targets (original paired images) compared to novels (never before seen images) in an explicit associative memory recall task. CA_3DG responded to lures (different exemplars of original images) as though they were novel, and mispaired images (originally paired with different cue names) as it would to targets, indicative of a task-neural discrimination profile. CA_1 activity, conversely, reflected task-sensitivity: mispaired images invoked responses similar to novels, while the response to lure images was similar to that of targets. These results show that a profile of mnemonic discrimination in CA_3DG and generalization in CA_1 (Bakker et al., 2008; Lacy et al., 2011) can be observed in explicit associative memory recall tasks. The dissociable responses between these subfields to task demands aligns with the

proposed role of the CA_1 within a larger reward/motivation system (Lisman & Grace, 2005).

CHAPTER 3

The Effects Of Reward Value On Item Representation In Human Hippocampus

How various motivations and drives affect the mechanisms supporting memory in humans is a topic that needs to be addressed in order to build a complete understanding of these mechanisms. Unlike rodents that have largely been the focus of research on the mechanisms of memory, humans' ability to recall and learn is heavily motivated by their own internal goals as well as the expressed demands of external forces (while animals are, of course, motivated by a number of internal and external forces, such as reward-seeking and pain aversion, these processes can not compare to the richness of humans' motivations). Furthermore, humans can, and do, employ their own strategic methods when tasked to remember materials (Benjamin, 2007). The human hippocampus exhibits a number of connections to both reward/salience regions (Gasbarri, Packard, Campana, & Pacitti, 1994; Lisman & Grace, 2005; Samson, Wu, Friedman, & Davis, 1990) and executive control areas (Preston & Eichenbaum, 2013), and how these systems modulate the medial temporal lobe (MTL) is a topic of great interest.

Rewards have been shown to modulate the mnemonic processes supported by the human MTL. Neuroimaging studies have demonstrated increased activity within the hippocampus and other MTL regions during encoding of high-value items relative to low-value items (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Wolosin, Zeithamova, & Preston, 2012). Furthermore, information about an item's reward value is represented by the distributed activation patterns of the hippocampus (Wolosin et al., 2013). However, it is unknown whether and how value modulates

specific item representations within the MTL. It is unclear whether the increase in activity for high-value items over low-value items reflects increased encoding of high-value items or a decrease in encoding of low-value items.

The effects of reward seen in the hippocampus indicate that reward somehow modulates the canonical subfield functions, such as pattern separation (the process of encoding overlapping information in such a way as to maximize differences and prevent catastrophic mnemonic interference (Kesner, 2013c)) and completion (the process of a recall based on degraded stimuli (Kesner, 2013b)). It is these two processes that are proposed as the driving forces behind successful memory encoding and retrieval. Emotional arousal as induced via salient images before an encoding session has been shown to increase pattern separation in a behavioral task (Segal, Stark, Kattan, Stark, & Yassa, 2012), and it is likely that reward will produce similar effects.

Several studies have shown that the combined CA_3/dentate gyrus (CA_3DG) (these regions can not be separated with current human imaging resolutions) exhibits activity reflective of pattern separation processes (Bakker et al., 2008; Lacy et al., 2011) although caution should be used when attributing discrimination responses to formal pattern separation. It has not been shown, however, that the CA_3DG response at encoding is indicative of later discrimination at recall, nor that reward has an effect on discrimination responses.

The present study addresses these questions by collecting high spatial resolution fMRI data during both encoding and retrieval in a reward-incentivized visual pattern separation memory task.

Hypotheses

- The similarity of brain activity patterns within MTL regions during presentation of face stimuli will be greater between those assigned the same reward value than between faces assigned different values.
- The within-value neural similarity for high value faces will be greater than that of low value faces.
- Greater reward values will increase the fidelity with which individual faces (and their constituent features) are stored in memory.
- Functional coupling between the reward regions, VTA and NAcc, will be greater at encoding for high value items.

Methods

Participants

Eighteen participants members of the University of California, Los Angeles (UCLA) community were recruited to participate in this study (15 females; mean age = 23.82 years, range 19 to 33 years). All participants were right-handed, native English speakers, had normal or corrected-to-normal vision and were free from any history of neurological or psychiatric symptoms. Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated for their time.

Experimental Design

Subjects studied photographs of faces (targets) in three blocks during fMRI scanning sessions. Half of the targets were cued with a high monetary value ($1.00) and half with a low value ($0.05), which subjects would earn for correct subsequent recognition. On each trial a monetary cue was presented first for 1 sec, followed by a

delay period during which a fixation cross was displayed on the screen. This fixation period was jittered between 3-6 sec to allow for increased ability to model the hemodynamic response of each event. The target image of a face was then presented for 3 sec. Following each learning image was a simple odd/even discrimination task. This task served both to distract subjects from rehearsing the target they just learned and as an active baseline wherein mind wandering is minimized as well as its associated hippocampal activity. A total of 66 faces were presented to subjects during the learning phase, 22 per block.

A recognition block followed each learning block. In this phase probe photographs were presented to subjects, who were asked to make old/new judgments. Probes consisted of similar photographs (view angle changed by 45 or 90 degrees) of

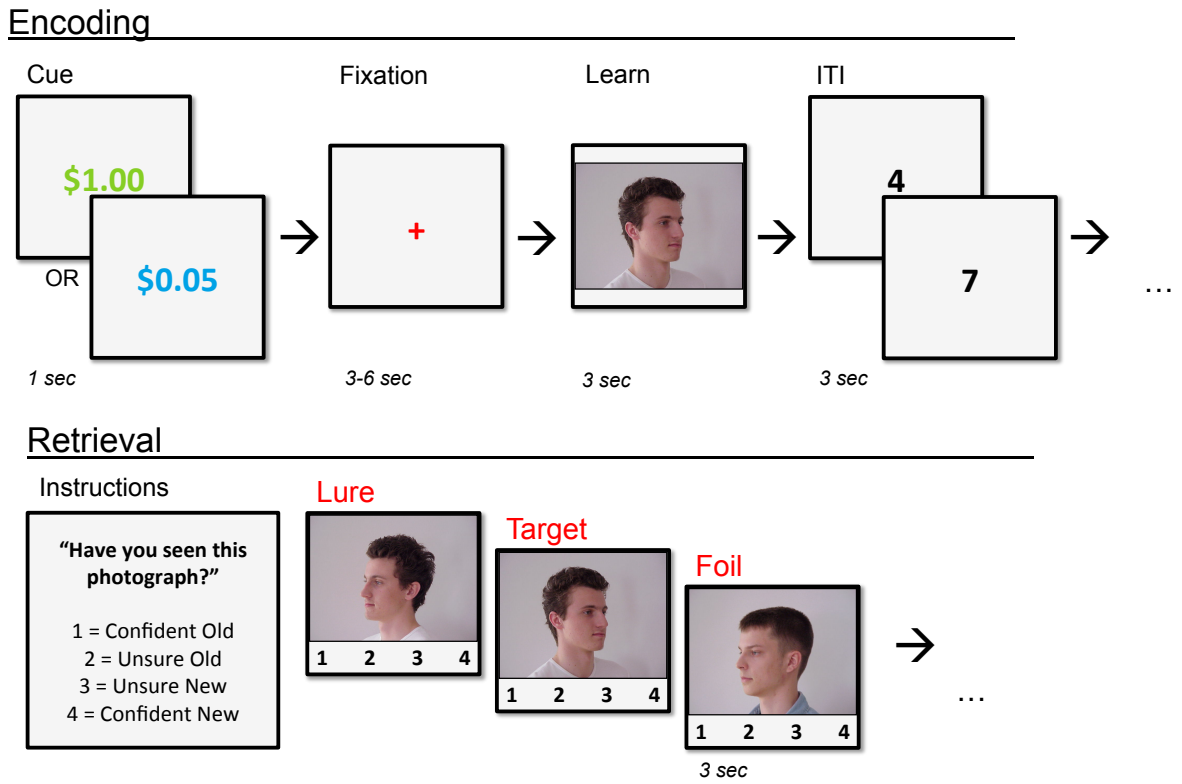


Figure 3.1: Experimental design.

the same face as every target learned (lure), re-presentations of the targets themselves, as well as completely new faces (foil). All targets, as well as a lure for each target, and six foil faces were presented in each block for a total of 50 probe faces per block. During recognition, subjects determined whether photographs were old (targets) or new (lures and foils) using a four-level scale (confident old, unsure old, unsure new, confident new). Subjects were penalized $0.15 for false alarms. Subjects were paid the amount earned during the retrieval phase at the end of the experiment. See Figure 3.1.

Scanning Parameters

Both phases (learning and recognition) were performed in the scanner, using high-resolution fMRI sequences. All scans were acquired on a 3T Siemens Tim Trio MRI system at the UCLA Staglin IMHRO Center for Cognitive Neuroscience. Functional data were collected using a T2*-weighted gradient-echo planar imaging (EPI) pulse sequence (repetition time (TR) = 3000 ms; echo time (ET) = 33 ms; 30 slices aligned perpendicular to the long axis of the hippocampus; 1.6 x 1.6 x 2.2 mm voxel size) using a GRAPPA acceleration factor of two to achieve the high spatial resolution. A T1-weighted, high-resolution sequence (magnetization-prepared rapid-acquisition gradient echo (MPRAGE), 240 sagittal slices, 1 x 1 x 1 mm voxels) was run to collect anatomical data of the whole brain. Additionally, a T2*-weighted high-resolution anatomical scan of the in-plane hippocampus structure was collected (30 slices aligned perpendicular to the long axis of the hippocampus, 0.4 x 0.4 x 2.2 mm voxels). This scan was used to create the anatomical regions of interests (ROIs) of the hippocampal subfields. Visual stimuli were presented via MRI-compatible LCD goggles. Subject responses were recorded with a four-button box.

ROI Creation, Registration and Pre-Processing

In order to extract activity from the hippocampal subfields, hand drawn ROIs were created for each subject using their high-res coronal structural image and following conventional region boundaries (Duvernoy et al., 2013; Frankó et al., 2014), Figure 3.2. Due to its small size, we included the CA_2 region in our combined CA_3DG ROI, which we henceforth refer to as $CA_{2,3}DG$. Anatomical ROIs were also created for the ventral tegmental area (VTA) and nucleus accumbens (NAcc) based on previous studies (Adcock et al., 2006). All fMRI BOLD data extraction analyses (i.e., measurements of the mean condition-specific activity level within each anatomical ROI) were performed in each subject's native EPI space, and several registration steps were taken to move the ROIs into this space. Registration was completed using Advanced Normalization Tools (ANTs) (Avants et al., 2010, 2011), which allowed for multiple transformations to be applied together such that images are only interpolated once. A template EPI space was created for each subject using ANTs based on exemplars image from each of the seven scanner runs. This template was used rather than a match bandwidth image as it better captures any small drifts in the FoV that might have resulted from subject motion over the course of the scanning session. Transformations were calculated for each scan run to this template, and the transform of the template to the subject's high-res hippocampal image was likewise calculated. The reverse transformations of the EPI-

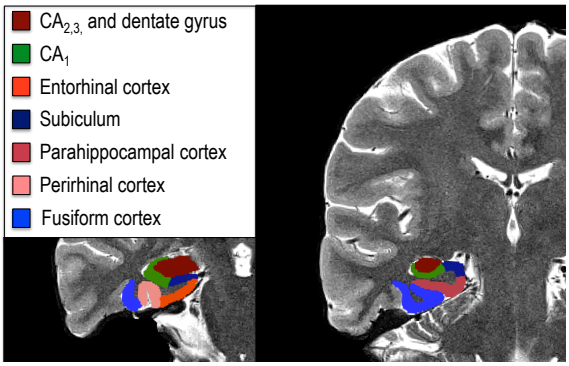


Figure 3.2: Medial Temporal Lobe Regions of interests. Example hand-drawn medial temporal lobe subfield ROIs (only left hemisphere shown) overlaid on an exemplar subject's high-resolution image.

template to high-res image and of the run EPI to EPI-template were applied to the ROIs. This series of transforms produced run-specific ROIs in EPI space. For group level analyses, additional registration steps were applied. Each subject's high-res hippocampal image was transformed into a study specific brain template, which was created with ANTs using all subjects' high-res images.

Several pre-processing steps were applied to prepare the data for statistical analysis. The data were corrected for motion during the scan run using FSL's (Jenkinson et al., 2012) FLIRT (Jenkinson et al., 2002) (FMRIB's motion correction linear image registration tool). Additionally, slice time correction was applied to the data. In an effort to prevent the blending of activity across neighboring hippocampal subfields, the data used in the ROI analyses were not spatially smoothed. However, when generating activity with reward area ROIs, a 5mm Gaussian smoothing kernel was applied to each subject's EPI data to reduce noise.

Analyses

Value encoding. One of the unique aspects of this data set is the acquisition of fMRI data during both encoding and retrieval. Having data from both task phases allows us to calculate encoding-retrieval similarities (ERS), a version of representational similarity analysis (RSA) that looks at the degree of correlation between the distributed pattern of activity within a region at different time points (Ritchey et al., 2013; Wing, Ritchey, & Cabeza, 2015). This is in contrast to traditional RSA analyses, which finds the correlation within a region between conditions (Kriegeskorte, 2008). Here we use both approaches to answer different questions. Neural similarities were quantified by calculating Pearson's correlations between voxel-by-voxel parameter estimates of each

trial event. These were created using a variant of the beta-series approach (Mumford, Turner, Ashby, & Poldrack, 2012; Rissman, Gazzaley, & D'Esposito, 2004), which models each individual trial's activity.

To determine the effects of value modulation of specific item representations within the MTL, the correlation between each target as seen during the learning phase (encoding) and when it was again presented during the recognition phase (retrieval) was calculated. The resulting ERS measures were then separately averaged for high-value and low-value items. The average correlation between all novel foil faces and each target face as presented during the learning phase was calculated as a baseline measure of the similarity of the processes engaged during encoding and recognition;

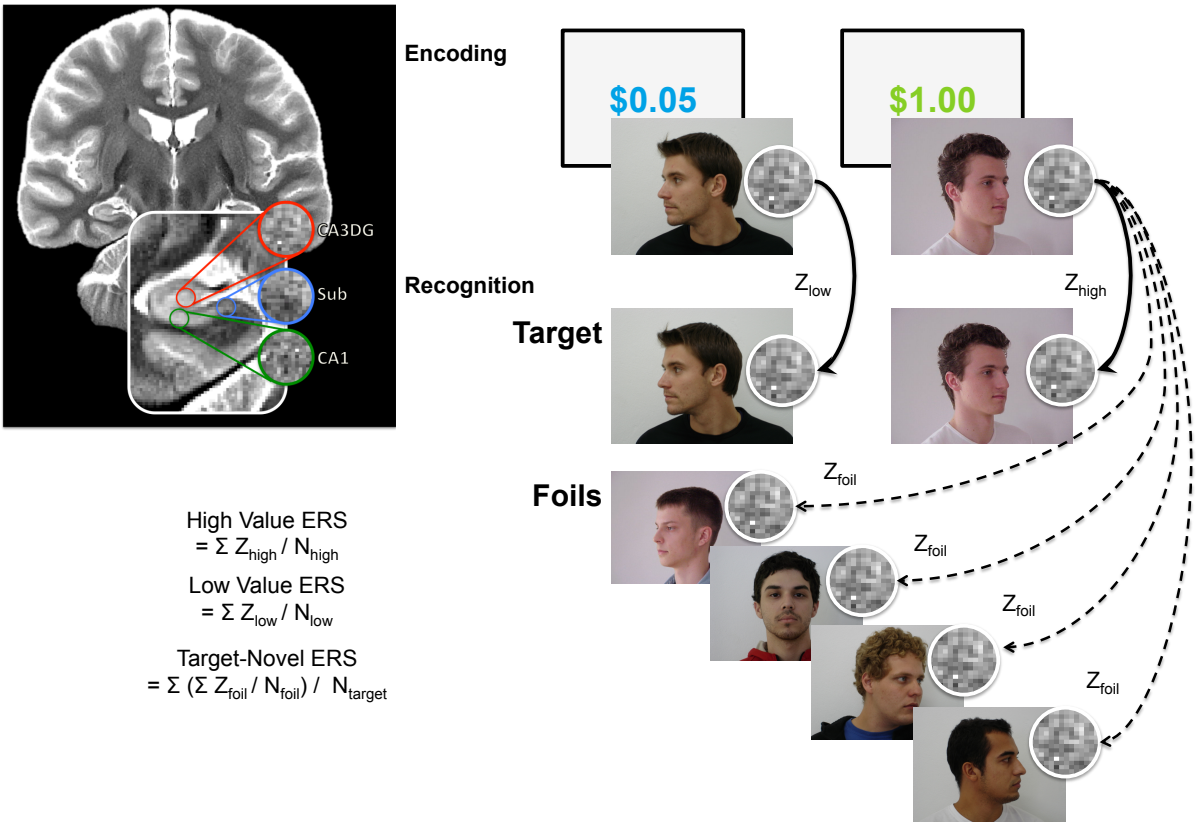


Figure 3.3: Calculation of Encoding Retrieval Similarity (ERS) Measures. ERS scores were separately averaged for each value level; correlations between each face at encoding and all novel foil stimuli per run were calculated for comparison purposes.

this allowed us to isolate item-specific representations (i.e. ERS effects linked to re-encountering the exact same photographs) from general engagement with and/or processing of the presented faces (Figure 3.3).

To replicate the findings of Wolosin et al. (2013) that there is a greater degree of similarity at encoding between pairs of same value items than between pairs of different-value items (Wolosin et al., 2013), we calculated within and between value similarities. For each individual face stimulus presented during encoding trials, its BOLD activity patterns estimates were correlated with those for all other faces. Within-value similarity estimates were created by taking the average correlation between all faces that shared the same value (High--High & Low--Low). Similarly, Between-value similarity is the average of correlation between faces that differed in their value (High--Low). We similarly calculated Within-value and Between-value similarities for activity during the re-presentation of target images during the retrieval scanner runs.

Reward regions connectivity with MTL. To hone in on the effects of reward within MTL regions we performed a psychophysiological interaction (PPI) analysis on encoding phase activity (O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). To see how activity of the VTA and NAcc modulate encoding in the hippocampus, we extracted the mean activity within our anatomically-defined VTA and NAcc ROIs from every volume (TR) of the preprocessed and spatially smoothed learning phase EPI scans. The interaction of these regions' activity and regressors modeling this phase was calculated with FSL's FEAT tools to create parameter estimates of functional coupling. Modeling of the encoding task is based on the value assigned at learning, and correct recognition at testing. For a target image at encoding to be considered as

subsequently remembered, that face at recognition needed to be correctly identified as old (hit) AND its 45 degree rotated lure image needed to be identified as new (correct rejection). If the target is a hit, but its lure is false alarmed to (called old), the target at learning would be considered subsequently familiar; if the target was mislabeled as new (miss), the target at learning would be considered subsequently forgotten regardless of the response to the lure.

Our model investigating activity during encoding phase consisted of five regressors, represented by 2-sec boxcars convolved with a double-gamma hemodynamic function: one each for high and low value for the cue presentation and one each for the image presentation binned based on subsequent recall (remembered, familiar, and miss). The ITI period remained unmodeled and served as an "active baseline" (Stark & Squire, 2001). Six standard motion parameters, as well as their first and second derivatives, were included in this model as regressors of no interest.

To be able to compare the relative coupling of reward and MTL regions between conditions we ran separate PPI models for each regressor. Each model included all five task regressors, an estimate of activity from the seed region, and the interaction between this activity and one of the five task regressors. These resulted in parameter estimates (betas) for each subject indexing the coupling of VTA and NAcc with all task conditions occurring during each run. Beta estimates calculated from these models are converted to percent signal change (Mumford, 2007) for ease of interpretation. These signal change estimates within each MTL ROI were extracted for conditions in each run, and a weighted average based on the number of trials per run was then calculated to generate across-run mean signal change estimates.

Results

Behavioral Data

To evaluate the effect of our value manipulation on subsequent recall of target faces we first examined the behavioral data. Overall, high-value faces were correctly recognized as old (*mean* = 71.80%, *SD* = 16.83%) more often than low-value faces (*mean* = 65.07%, *SD* = 15.07%; $t_{(17)}$ = 2.80, *p* = 0.012). However, when we further divide recall trials by subjects' responses to both target repeats and their lures, the proportion of remembered high value faces (*mean* = 43.52%, *SD* = 16.71%) was not significantly greater than low value faces (*mean* = 40.07%, *SD* = 18.29%; $t_{(17)}$ = 1.04, *p* = 0.314) (Figure 3.4). While reward values seems to have influenced recognition of old faces, correct rejection rates for lure images did not vary between high and low reward values assigned to the associated target image (high: *mean* = 61.95%, *SD* = 18.36%; low: *mean* = 58.50%, *SD* = 18.36%; $t_{(17)}$ = 0.98, *p* = 0.338).

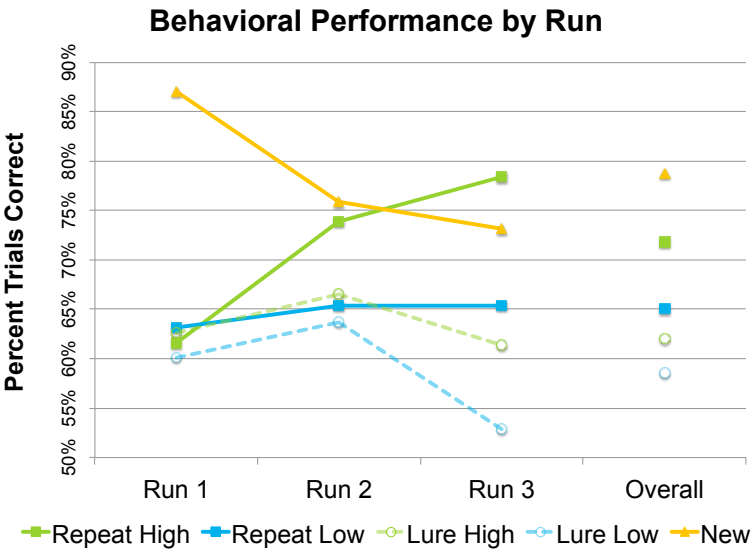


Figure 3.4: Behavioral Performance by Run.

Value Representation at Encoding

In contrast to the results of Wolosin et al., (2013) no MTL region showed greater overall similarities between within-value faces that between-value faces (all $p > 0.11$) (see appendix chart). However, within-value similarity of high-value faces (High--High) within the right fusiform cortex (*mean* = 0.012, *SEM* = 0.003) is significantly greater than low-value faces (Low--Low) (*mean* = 0.006, *SEM* = 0.003; $t_{(17)} = 3.83$, $p = 0.001$). In this

Table 3.1: Representational Similarity Estimates for Reward Values at Encoding and Retrieval

Region		High-Low±SEM	$t_{(17)}$	High-Between±SEM	$t_{(17)}$	Low-Between±SEM	$t_{(17)}$
				Encoding			
CA1	Right	0.0006±0.0027	0.2398 ns	-0.0003±0.0021	0.1613 ns	-0.0010±0.0014	0.6843 ns
	Left	0.0024±0.0029	0.8116 ns	0.0006±0.0018	0.3397 ns	-0.0018±0.0022	0.8164 ns
CA23DG	Right	0.0007±0.0027	0.2499 ns	0.0018±0.0017	1.0404 ns	0.0011±0.0029	0.3756 ns
	Left	-0.0012±0.0023	0.5140 ns	-0.0022±0.0024	0.9281 ns	-0.0010±0.0015	0.6791 ns
Sub	Right	-0.0035±0.0031	1.1192 ns	-0.0009±0.0025	0.3660 ns	0.0026±0.0017	1.5234 ns
	Left	-0.0019±0.0030	0.6399 ns	-0.0010±0.0019	0.5265 ns	0.0009±0.0019	0.4704 ns
ERc	Right	-0.0020±0.0054	0.3616 ns	0.0003±0.0044	0.0618 ns	0.0022±0.0030	0.7378 ns
	Left	0.0003±0.0038	0.0743 ns	-0.0023±0.0029	0.7911 ns	-0.0026±0.0020	1.2744 ns
PHc	Right	0.0003±0.0020	0.1282 ns	-0.0011±0.0014	0.7886 ns	-0.0014±0.0018	0.7661 ns
	Left	-0.0017±0.0022	0.7910 ns	-0.0018±0.0019	0.9974 ns	-0.0001±0.0016	0.0644 ns
PRc	Right	0.0026±0.0037	0.7082 ns	0.0040±0.0029	1.3709 ns	0.0014±0.0019	0.7278 ns
	Left	-0.0004±0.0017	0.2486 ns	-0.0009±0.0014	0.6280 ns	-0.0005±0.0012	0.4181 ns
FUSc	Right	0.0063±0.0016	3.8260*	0.0030±0.0010	3.0546*	-0.0032±0.0016	2.0676~
	Left	0.0011±0.0031	0.3448 ns	-0.0001±0.0014	0.0667 ns	-0.0012±0.0021	0.5436 ns
				Retrieval§			
CA1	Right	0.0017±0.0051	0.3369 ns	0.0037±0.0024	1.5236 ns	0.0020±0.0033	0.6095 ns
	Left	0.0044±0.0036	1.2165 ns	0.0022±0.0020	1.0945 ns	-0.0023±0.0032	0.7039 ns
CA23DG	Right	0.0021±0.0041	0.5176 ns	0.0008±0.0028	0.2783 ns	-0.0013±0.0018	0.7350 ns
	Left	0.0028±0.0047	0.5909 ns	0.0029±0.0019	1.4849 ns	0.0001±0.0032	0.0347 ns
Sub	Right	0.0034±0.0043	0.8054 ns	0.0015±0.0024	0.6284 ns	-0.0019±0.0022	0.8548 ns
	Left	0.0058±0.0036	1.6081 ns	0.0033±0.0027	1.1940 ns	-0.0025±0.0014	1.7488 ns
FUSc	Right	0.0020±0.0028	0.7266 ns	0.0019±0.0012	1.5886 ns	-0.0002±0.0022	0.0829 ns
	Left	-0.0007±0.0034	0.2229 ns	-0.0013±0.0023	0.5780 ns	-0.0006±0.0015	0.3974 ns

§ Retrieval RSA values only calculated for regions previously displaying significant ERS or RSA results; * $p < 0.05$; ~ $p < 0.1$

region the Within-high value similarity is greater than the baseline between-value similarity (High--Low) (*mean* = 0.009, *SEM* = 0.003; $t_{(17)}$ = 3.05, *p* = 0.007). Although the Within-low value similarity is less than the baseline between-value, this contrast just fails to pass significance ($t_{(17)}$ = 2.07, *p* = 0.054). No other MTL regions showed any differences resulting from value (all *p* > 0.28) (see Table 3.1).

Encoding-Retrieval Similarity by Value

Although we did not find any effect on similarities based on values within the encoding phase, several hippocampal regions showed value modulated differences in item target-to-target correlations between encoding and retrieval phases. Within the right hippocampus, the ERS for high-value targets (*mean* = 0.011, *SEM* = 0.004) was significantly greater than low-value targets (*mean* = -0.003, *SEM* = 0.003; $t_{(17)}$ = 2.56, *p* = 0.020). The left hippocampus showed a nearly significant trend in the same direction (high: *mean* = 0.012, *SEM* = 0.004; low: *mean* = -0.004, *SEM* = 0.005; $t_{(17)}$ = 2.09, *p* = 0.051). Compared to the baseline of encoding--novel similarity (*mean* = 0.005, *SEM* = 0.002), the value-related modulation of ERS in the right hippocampus represents a decrease for low value ($t_{(17)}$ = 2.15, *p* = 0.046) rather than an increase for high value ($t_{(17)}$ = 0.84, *p* = 0.413). A similar pattern of results was observed in left hippocampus, albeit not quite significant (encoding--novel similarity: *mean* = 0.010, *SEM* = 0.005; high value: $t_{(17)}$ = 0.78, *p* = 0.448; low value: $t_{(17)}$ = 2.03, *p* = 0.057).

When evaluating these finding at the subfield level, only the right $CA_{2,3}DG$ showed a significant difference between similarities for high (*mean* = 0.008, *SEM* = 0.005) and low values (*mean* = -0.007, *SEM* = 0.004; $t_{(17)}$ = 2.22, *p* = 0.041), although the right CA_1 (high: *mean* = 0.014, *SEM* = 0.006; low: *mean* = -0.003, *SEM* = 0.007; $t_{(17)}$

= 1.99, p = 0.062), left $CA_{2,3}DG$ (high: *mean* = 0.008, *SEM* = 0.006; low: *mean* = -0.006, *SEM* = 0.005; $t_{(17)}$ = 1.93, p = 0.070), left subiculum (high: *mean* = 0.017, *SEM* = 0.004; low: *mean* = -0.002, *SEM* = 0.009; $t_{(17)}$ = 1.81, p = 0.087),and right fusiform cortex (high: *mean* = 0.029, *SEM* = 0.006; low: *mean* = 0.013, *SEM* = 0.006; $t_{(17)}$ = 2.07, p = 0.053) all show strong trends in the same direction. As in the whole hippocampi, the difference in ERS based on value is driven largely by a decrease in similarity for low value faces. Low value faces showed a reduced correlation between encoding and retrieval phases when relative to that of encoding--novel in the both left (encoding--novel: *mean* = 0.011, *SEM* = 0.005; $t_{(17)}$ = 2.69, p = 0.015) and right $CA_{2,3}DG$s (*mean* = 0.003, *SEM* = 0.002; $t_{(17)}$ = 2.15, p = 0.046), and in the right fusiform (*mean* = 0.020, *SEM* = 0.004; $t_{(17)}$ = 2.15, p = 0.045); high value ERS did not differ significantly from baseline encoding--novel ERS in any of these region (L $CA_{2,3}DG$: $t_{(17)}$ = 0.15, p = 0.883; R $CA_{2,3}DG$: $t_{(17)}$ = 0.84, p = 0.413; R Fus: $t_{(17)}$ = 1.67, p = 0.112), although in the left subiculum did show a trend in this direction (*mean* = 0.005, *SEM* = 0.004; $t_{(17)}$ = 2.03, p = 0.058), while low value didn't differ from baseline ($t_{(17)}$ = 1.00, p = 0.331). Only in the right CA_1 did the difference between high and low value not result from a significant decrease for one value only, rather the encoding--novel ERS feel between the two value levels, and did not differ from either (*mean* = 0.006, *SEM* = 0.003; high: $t_{(17)}$ = 1.52, p = 0.148 ; low: $t_{(17)}$ = 1.53, p = 0.145) (see Table 3.2).

These results indicate that value does not increase the specific representation of the faces per se, rather the difference may reflect less consistent encoding or engagement with the low value faces. This is further supported by the observation that, although not significant, in the right $CA_{2,3}DG$ the similarity between each learned target

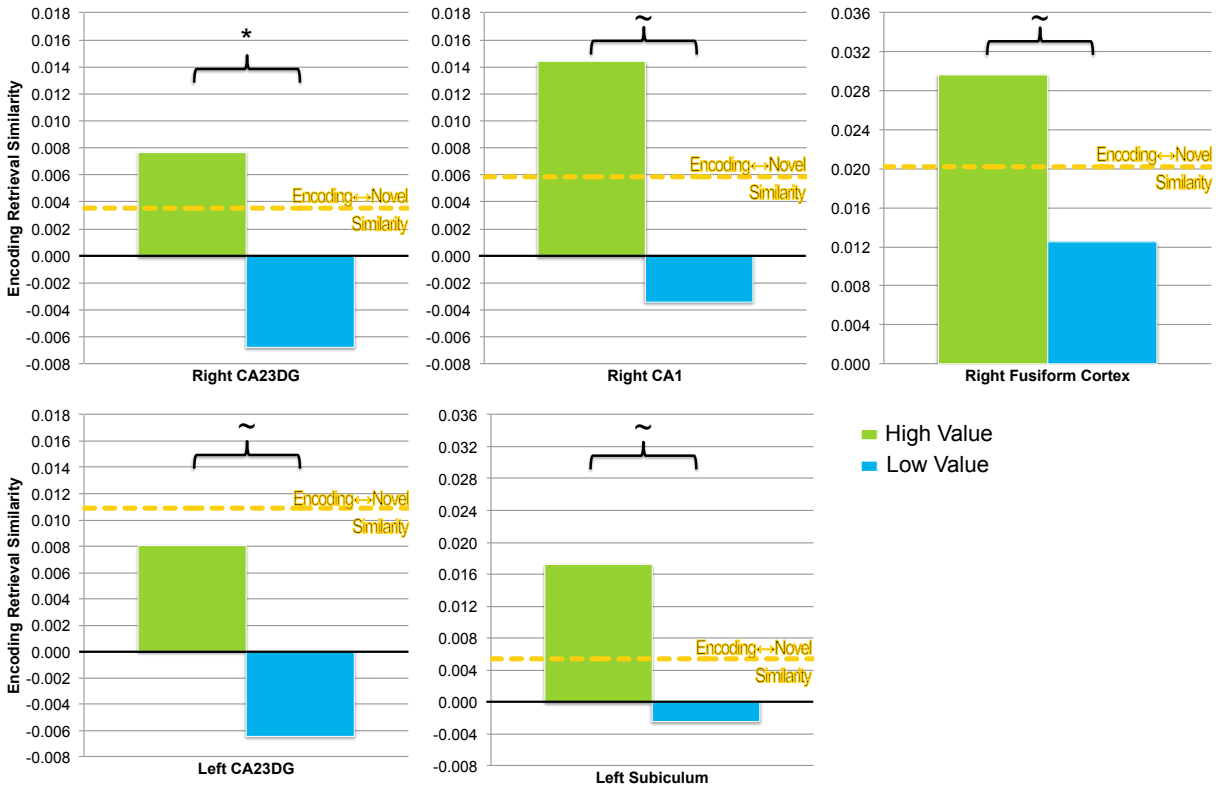


Figure 3.5: Encoding-Retrieval Similarity by Reward Value. Only regions with significant, or trending, difference between high and low ERS shown, see Table 3.2 for full results.

and each other probed target of the same value (excluding its own matching target probe) showed a similar profile of greater ERS values for high value (*mean* = 0.012, *SEM* = 0.007) over low value (*mean* = -0.001, *SEM* = 0.005), at comparable levels of magnitude ($t_{(17)}$ = 1.12, p = 0.275), reflecting that the general value class of the faces is at least partially driving our ERS effects.

Effective Connectivity Between Reward Regions and MTL During Encoding

To examine the modulating influences that the reward system has on MTL regions we next compared the functionally coupling between these regions and the VTA and NAcc. Activity within the left CA_1 region reflected greater coupling with the VTA during the presentation of high-value cues (*mean* = 0.0010, *SEM* = 0.001) than low-value cues (*mean* = -0.0016, *SEM* = 0.001; $t_{(16)}$ = 2.21, p = 0.042). Conversely, within

the right fusiform the opposite pattern was found: high-value cues (*mean* = -0.002, *SEM* = 0.001) showed a trend towards less coupling with the VTA than low-value cues (*mean* = 0.0007, *SEM* = 0.001; $t_{(16)}$ = 2.03, p = 0.059). No other region showed differential functional coupling for cue presentations between VTA or NAcc as a function of assigned reward value (all $p > 0.13$). These results would indicate the differences seen in the RSA and ERS results are likely not due to direct modulation by these regions, as none of the regions that showed a benefit in either of these representation analyses for high-value items showed

Value Representation at Retrieval

Unlike during the encoding phase, no MTL regions showed any value differences between within-value similarities of target images (all $p > 0.13$). In the right fusiform, the only region to show significant value representation modulation, high within-value faces

Table 3.2: Encoding-Retrieval Similarity by Reward Value Within MTL

Region		High-Low±SEM	$t_{(17)}$	High-New±SEM	$t_{(17)}$	Low-New±SEM	$t_{(17)}$
CA1	Right	0.0178±0.0089	1.9948~	0.0084±0.0055	1.5160 ns	-0.0119±0.0078	1.5267 ns
	Left	0.0086±0.0097	0.8810ns	0.0025±0.0059	0.4280 ns	-0.0068±0.0070	0.9716 ns
CA23DG	Right	0.0144±0.0065	2.2174*	0.0041±0.0049	0.8391 ns	-0.0104±0.0048	2.1501*
	Left	0.0145±0.0075	1.9336~	-0.0009±0.0059	0.1491 ns	-0.0169±0.0062	2.6911*
Sub	Right	-0.0012±0.0082	0.1429 ns	0.0039±0.0072	0.5363 ns	0.0036±0.0045	0.8025 ns
	Left	0.0197±0.0109	1.8155~	0.0124±0.0061	2.0312~	-0.0106±0.0106	0.9992 ns
ERc	Right	0.0105±0.0104	1.0084 ns	0.0004±0.0088	0.0424 ns	-0.0139±0.0102	1.3644 ns
	Left	0.0159±0.0110	1.4448 ns	0.0033±0.0132	0.2532 ns	-0.0117±0.0094	1.2519 ns
PHc	Right	0.0053±0.0073	0.7334 ns	0.0018±0.0065	0.2834 ns	-0.0058±0.0036	1.6027 ns
	Left	0.0032±0.0079	0.4051 ns	-0.0050±0.0071	0.7031 ns	-0.0077±0.0081	0.9527 ns
PRc	Right	-0.0020±0.0088	0.2272 ns	-0.0011±0.0071	0.1489 ns	0.0000±0.0059	0.0030 ns
	Left	0.0051±0.0052	0.3391 ns	-0.0034±0.0064	0.6039 ns	-0.0086±0.0060	0.1662 ns
FUSc	Right	0.0171±0.0083	2.0742~	0.0100±0.0059	1.6724 ns	-0.0099±0.0046	2.1559*
	Left	0.0028±0.0051	0.5463 ns	-0.0028±0.0064	0.4352 ns	-0.0062±0.0057	1.0884 ns

* $p < 0.05$; ~ $p < 0.1$

(*mean* = 0.018, *SEM* = 0.005) were very slightly more similar than low within-value faces (*mean* = 0.014, *SEM* = 0.004), but this difference did not approach significance ($t_{(17)}$ = 9.95, p = 0.357)) (see Table 3.1). This lack of reinstatement of value information within the MTL during retrieval aligns with the results of our ERS analyses that indicate that previous findings of value representation within the MTL may, in fact, reflect different encoding processes, and thus would not be expected at retrieval.

Discussion

Our findings indicate that items assigned a low reward value are putatively processed within the medial temporal lobe, particularly the right $CA_{2,3}DG$, in a less consistent manner than higher value items at encoding. While high value faces showed greater similarity between encoding and retrieval presentations than low value faces, the representation of those high value items were not more similar between phases. When compared to the baseline similarity between the representation of target faces at encoding and novel faces at retrieval, the similarity between target faces assigned low values at encoding and the same face presented again during retrieval was reduced. No MTL region showed greater than baseline between phase similarities for high value items. While the decrease for low value items relative to baseline was seen in several regions in both hemispheres of the MTL, only in the right $CA_{2,3}DG$ was the difference between high and low value significant. The sensitivity of $CA_{2,3}DG$ to differences in encoding processes is in line with the proposed role of CA_3 as the region responsible for rich pattern-separated representations of mnemonic information (Kesner, 2013b). The lateralization of our effect to the right $CA_{2,3}DG$ likely reflects the importance of spatial information, particularly viewpoint (Lambrey et al., 2008), in our discrimination task.

This finding is in contrast with our hypothesis that high value rewards would increase the fidelity of the neural representation of an item and its reward value. The differentiation between motivational reward context in the representational similarity patterns of the hippocampus seen in previous work (Wolosin et al., 2013) was theorized to result from the binding of reward context with its associated event. We hypothesized that this binding of reward context would result in greater reactivation of reward information at retrieval, and thus a greater encoding-retrieval similarity. The lack of difference between the ERS values of high value items and the similarity between studied items and novel, unrewarded, items refutes this hypothesis. At retrieval, MTL regions did not show strong reactivation of reward representation for high value items; neither did high reward values result in a more distinct representation of the face itself within these regions. Either of these factors would have increased the similarity of retrieval presentation of target images with the activity observed at its original presentation.

The relative decrease in encoding-retrieval similarity for low value items might have resulted from a number of different underlying sources. While, as noted above, it had been theorized that differences in representational similarities between assigned values resulted from a representation of that value encoded along with its associated context, an alternative framework suggested for interpreting this results is that these differences are due to encoding strategies that differed between reward conditions. Our results would seem to indicate this framework. As the task demands differed between the two phases, encoding verse memory evaluation, any similarity between these phases cannot be attributed to any one computation and must instead result from the

coherence of items and/or value representation. The similarity between the presentation of target faces at encoding and novel foil images at retrieval lack any such coherence in item or value, and thus can be viewed as a baseline measure of a region's engagement with our stimuli. That low value items show lower ERS estimates than this baseline would indicate different engagement with these faces. Items assigned a low reward value are putatively processed in a less consistent manner than higher value items. We did not collect reports of strategies employed by participants, so we cannot say for certain this apparent difference in encoding processes resulted from experientially divergent engagement. Nonetheless, our findings would seem to indicate that encoding processes, not stimulus representation, differentiates medial temporal lobe activity due to value at encoding.

In further support of the theory that reward values result in different processes within MTL regions rather than different representations, we did not find any differences in the representational similarities of high and low value target images at retrieval. The lack of reinstatement of reward context during recall indicates that this context was not strongly bound to the item memory at the time of encoding. This lack of representation of the value itself within the hippocampus supports the hypothesis that this region is 'representationally agonistic' (Huffman & Stark, 2014). In a memory system a representationally agonistic region is one that does not itself hold any of the categorical or conditional content of memories. It has been theorized that medial temporal lobe cortices, but not the hippocampus itself, hold and represent at least some of the content of memories, while the hippocampus is responsible for more computational processing of memory, in a content agnostic manner (Azab et al., 2014). Interpreted within this

model, the inability to differentiate between the representational similarities of reward context in any hippocampal region is expected.

Despite not showing any increase over baseline, our ERS result does not rule out the possibility that deeper processing is occurring during encoding of high value items. Indeed, the increase in RSA values within the right fusiform for high value items over between value items at encoding might suggest more consistent, and hence more focused processing of these face stimuli. This conforms to the results of previous studies that have shown increased MTL activity for high value items (Adcock et al., 2006; Wolosin et al., 2012). Alternatively, assigning a face to a high value might increase overall focus and reduce extraneous processing. This reduction, rather than the focusing of processing, may have driven our finding of greater within-value similarities for the high-value faces. Regardless of the causal mechanism, our RSA results would indicate that assigning a face stimulus to a high reward value benefits its processing in the right fusiform cortex. The improvement of processing within this face sensitive region, and no other MTL region, might explain the recognition, but not remember, benefit for high value faces over low value. Extra-hippocampal MTL cortex has been implicated as the regions responsible for familiarity-based item recognition, while the hippocampus is thought to be required for the more involved experience of contextual remembering (Ranganath et al., 2004).

An interesting result of our psychophysiological interaction analysis is that almost none of the regions that showed differential ERS or RSA responses based on assigned reward value displayed differential functional coupling with reward regions. Only the left CA_1 showed greater coupling with the VTA for high value cues than low value ones, and

no regions showed differential coupling with the NAcc. The overall lack of difference in PPI values within the right fusiform would indicate that the increase in within-value RSA noted above likely did not result from direct influences of either the VTA or NAcc. Instead, one or both of these regions may have contributed indirectly, potentially via the prefrontal cortex (Cohen, Rissman, Suthana, Castel, & Knowlton, 2014), to encoding processing within the MTL. Unfortunately, the reduced field of view needed to acquire high spatial resolution MTL data precluded our ability to collect functional data from many other potentially relevant brain regions. Thus we were unable to characterize the larger-scale functional connectivity profile of midbrain and ventral striatal reward-related regions. Finally, the coupling between the CA_1 and VTA aligns with the findings of Chapter 2, which showed that this hippocampal subfield was uniquely sensitive to task demands. The CA_1 is thought to be a key node in a network connecting the hippocampus and the VTA (Lisman & Grace, 2005) and receives direct projections from that area (Gasbarri et al., 1994). Our results here further support the assertion that the CA_1 is the hippocampal region sensitive to situational motivations.

Conclusions

Here we show that the reward value that is assigned to a given face during encoding serves to modulate its neural similarity with other faces that share the same value within the right fusiform cortex. Furthermore, low value faces show decreased encoding-retrieval similarity in several hippocampal regions, particularly the right $CA_{2,3}DG$. These findings indicate that items assigned a low reward value are processed in a less consistent manner than higher value items in the hippocampus at encoding. The implication of our results here is that previous findings of value representation

within the MTL may, in fact, reflect different encoding processes, wherein low-value items less strongly and/or less consistently engage effective mechanisms for encoding stimulus-specific representations. The sensitivity of $CA_{2,3}DG$ to these differences in encoding processes is in line with the proposed role of the CA_3 as the hippocampal region responsible for rich pattern-separated representations of mnemonic information.

CHAPTER 4

Stimulus Driven Shift between Memory Retrieval and Encoding: Encoding Benefits via Mismatch Detection

One of the most fascinating, and computationally complex aspects of memory is the flexible shift between recalling of what has come before, and encoding what is happening now. This shift occurs constantly as we go about our lives. Both these mnemonic computations rely on the hippocampus, with different functional components assigned to the various subfields. The dentate gyrus (DG) subfield is often ascribed the role of pattern separation, a process of encoding memories whereby differences between similar memories are enhanced such that there is minimal overlap in their neural representation (Kesner, 2013a). Conversely, the CA_3 subfield of the hippocampus is thought to be involved in representing associations between event details and subsequently triggering the reactivation of these associations during cued retrieval, pattern completion (Kesner, 2013b). CA_3 responses will switch from those reflecting completion to those reflecting separation when the change in a cueing stimuli is different enough from what was previously encountered (S. Leutgeb & Leutgeb, 2007; Vazdarjanova & Guzowski, 2004). Several high-resolution fMRI (hr-fMRI) (Carr et al., 2010) studies have shown evidence for the presence of similar computations in humans (Bakker et al., 2008; Lacy et al., 2011; Yassa & Stark, 2011).

The CA_1 subfield of the hippocampus is thought to compare the retrieved mnemonic content of CA_3 with incoming signals from the entorhinal cortex. This process can be thought of as a comparison of memory-based predictions with emerging sensory reality (Kumaran & Maguire, 2007b). Correspondingly CA_1 has been shown to have

greater activity when one's memory-based predictions are violated, a putative prediction error signal (Chen et al., 2015, 2011; Duncan et al., 2012). This CA_1 prediction error signal is thought to be an important trigger for new learning – orienting an individual to attend to environmental cues in an effort to understand why the world failed to match expectations (Lisman & Grace, 2005). At baseline, when one is not actively trying to remember or recall, this is an effective system for minimizing the need to be encoding continuously. Studies have shown, however, that when presented with a task explicitly evaluating mnemonic probes, the response in CA_1 shifts from increased activity for mismatches to an increase for matches (Kumaran & Maguire, 2007a), reflecting the sensitivity of this region to task goals.

Despite this mechanistic framework for understanding mnemonic encoding and associative retrieval, only one study has shown empirical evidence that items that deviate from a previously encountered mnemonic context are later recalled to a greater degree than if they had not triggered the putative error signal of the CA_1. As a follow-up to their fMRI study, Chen et al. (2015) conducted a surprise recognition test for items that previously repeated or violated their first presentation order. This recognition test revealed a small, but significant, mismatch benefit, manifest as an increased hit rate for order violation items (Chen et al., 2015). Given the singularity of this result, and the weak magnitude of the effect, further experiments replicating this mismatch benefit are required.

That distinct, or isolated, items are more strongly recalled than other items has long been known (von Restorff, 1933), a perceptual saliency effect known as the von Restorff effect (Hunt, 1995). Furthermore, this von Restorff effect seems to necessitate

a properly functioning hippocampus (Kishiyama, Yonelinas, & Lazzara, 2004) and substantia nigra (Schomaker et al., 2014), consistent with the functional loop via the mesolimibic dopaminergic system (Lisman & Grace, 2005) regulating the encoding of long-term memories (Shohamy & Adcock, 2010) as a function of novelty. To the best of our knowledge, no prior study has used a mnemonic mismatch as the distinct aspect of a task designed to invoke a novelty detection response and trigger greater recall. The lack of evidence for a variant of the von Restorff effect that is driven exclusively by memory-based prediction error represents a significant gap in our understanding of the function the CA_1 subfield and the validity of the theory that this region serves as a comparator. Here we conduct a number of experiments that manipulate several factors relating to the presentation of mnemonic mismatches. Each experiment is followed by a surprise memory test in order to assess the effect of mismatch on subsequent memory performance.

Hypotheses

- Recall for items presented in a mismatching condition will be greater than those presented in a non-mismatching condition.
- Item memory for mismatching items presented in a well learned context will be greater than for those in less well learned contexts.
- When a mismatch occurs after the presentation of more items of a learned sequence, subsequent recall will be greater than if less of the sequence precedes the mismatch.

Table 4.1: Overview of experimental parameters for Experiments 1-8

	Conditions	Mismatch type	Days	Sequence length	Learning blocks	Mismatch position	Relationship judgments timing	# probes	Test
1	1. Mismatch 2. Repeat 3. New	Order Swap	1	4	3	3 & 4	End of sequence	4	Order
2	1. Mismatch 2. Repeat 3. New	Order Swap	1	4	3	3 & 4	End of sequence	4	Item and Order
3	1. Mismatch 2. Repeat 3. New	Order Swap	2	4	3	3 & 4	End of sequence	4	Item
4	1. Mismatch 2. Repeat 3. Scrambled	Novel item	2	4	3	3 or 4	End of sequence	4	Item
5	1. Mismatch 2. Repeat 3. Scrambled	Novel item	2	4	1 or 4	4	End of sequence	4	Item
6	1. Mismatch 2. Repeat 3. Scrambled	Novel item	2	5	3	4 or 5	• Learn - End of sequence • Mismatch – Semantic judgment each item	5	Item
7	1. Mismatch 2. Repeat 3. Scrambled	Novel item & Response Judgment	2	4	3	4	After each item	1	Item
8	1. Mismatch 2. Repeat 3. Shuffled	Novel item	2	4	3	4	End of sequence	4	Item

Experiment 1: Sequence Encoding and Reconstruction

Design

This task probes subjects' memory for the order of presented items. The task is comprised of three phases, all conducted on the same day: a learning phase, mismatch learning phase, and testing phase. Before the test phase subjects do not know that they are participating in a memory study; rather they are told that they are taking part in a study of how people make judgments about the relationships between items. In this way, learning of stimuli and their presentation order and any mismatch detection is incidental rather than intentional. All subsequent versions of the task (Experiments 2-8)

follow this three-phase structure (although the timing between phases varies), and all attempt to keep the memory test a surprise.

During the initial learning phase, subjects are presented with sequences of four items appearing one after another, for 1.25 sec each, with a brief 0.25-sec fixation cross between images. For each series, subjects determine how many of the items in that series met a specified relationship with the item that preceded it. For example, one relationship is “larger”, so subjects need to determine how many items in a series are larger than the item that preceded it. If the sequence is house, pencil, dog, bear, in that order, the number of larger images would be "2", as a dog is larger than a pencil, and a bear is larger than a dog, but a pencil is not larger than a house. Subjects are prompted to make their response, ranging from 0-3, times the relationship was met in the 2 seconds following the last item of the sequence. This task allows for strong encoding of order in an incidental manner.

This learning phase is broken up into three blocks and each block asks subjects to make judgments regarding different relationship including: "larger than the previous item," "would be seen with the previous item,” “made of the same material/stuff as the previous item.” All 90 sequences, presented in intact order, occurred during each block, for a total of three learning opportunities.

The mismatch phase immediately follows the learning phase. Two tasks make up this phase: an N-back task and the mismatch task. The N-back task sequentially presents images, and asks participants to indicate when an image is presented twice in a row (1-back). This task pre-exposes all of the images to be presented as new sequences during the ensuing mismatch phase. During the mismatch phase, which

strongly resembles the learning phase, subjects again see a series of four images, and again make judgments about the relationship between the images, namely: “Smaller than the previous item.” Critically, this phase presents three different types of sequences: 1) previously viewed intact sequences (“repeats”), 2) previously viewed sequences where the order of the final two items is swapped (“mismatch”), and 3) “new” sequences, comprised of the novel images from the N-back task. In total, 50 mismatch, 40 repeat, and 30 new sequences are presented.

Finally, after subjects have completed the mismatch phase they are given a surprise memory test for the order of the sequences they saw during the previous phase. Subjects are cued with the first item in a sequence and are presented with four images (the three images that make up the rest of the sequences and one randomly selected foil image from another sequence) and are prompted to rebuild the remaining sequence from those images. Subjects use a computer mouse to drag and rearrange these images, as needed, until they are satisfied with their response. After confirming that they are satisfied with the sequence they rebuilt, subjects are asked to state if the order they saw (and hopefully recreated) in the previous phase was the same as the one they saw before in the first phase, a different order, or a new sequence that was not previously shown. Figure 4.3 shows the design for this experiment.

Participants

Twenty (14 female) subjects completed the first experiment ($mean_{age}$ = 19.89 years). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time.

This data set excluded two subjects that performed below chance on rebuilding of all sequences, yielding a total data set of 18 (14 female) subjects.

Results

Experiment 1 was initially conducted as a pilot study before moving into the scanner, therefore no comparisons were planned and only trial counts were analyzed. Overall, subjects correctly rebuilt the sequences as seen in the mismatch phase 42.45% (*SD* = 10.96%). When broken down by trial type, repeat sequences were rebuilt correctly 68.09% (*SD* = 16.41%; 31.91% incorrect) of the time, novels were correctly rebuilt for 29.61% of sequences (*SD* = 12.85%; 70.39% incorrect), while mismatched sequences were rebuilt in the correct mismatch sequence order 29.65% (*SD* = 11.24%) of the time and in the original presentation sequence order 37.76% (*SD* = 12.12%), thus subjects placed items in a previously seen order in 67.41% of trials (32.59% were placed in a non-presented order), Figure 4.1. Chance rates for rejecting the lure image and rebuilding sequences in the correct order is one in twenty-four or about 4.2%, and the chance rate for rebuilding mismatch sequences in one of the two previously seen orders is one in twelve or about 8.3%, so the results here indicate performances well above chance.

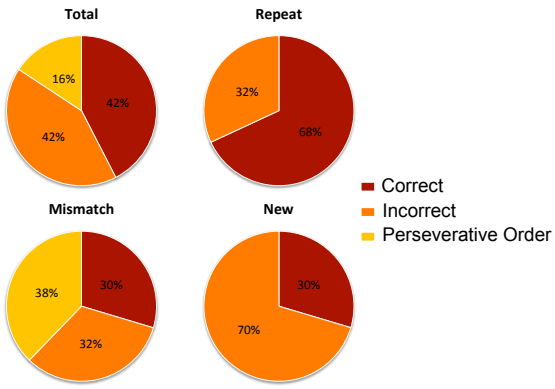


Figure 4.1: Results for Experiment 1. Only sequence rebuilding was tested.

Discussion

The results of this first experiment, while would allow for sufficient trial counts for analysis of scan data, do not show any

benefit of later recall for the presentation of a mismatch. This is not surprising as it is highly unlikely that any such benefit would be enough to supplant the learned orders seen over multiple previous exposures. However, it is clear that this incidental task design leads to fairly robust encoding of sequences, to the extent that participants could correctly reject lure items and rebuild the sequences on upward of 40% of all trials, and up to 68% of repeated sequences.

In order to detect if the expected mismatch benefit could be observed in the recall of the items that made up each sequence, rather than the order of the sequence, we next conducted an experiment that included a test for subjects' item memory into the design used in this experiment.

Experiment 2: Sequence Reconstruction and Item Test

Design

In this second version of the task, the learning and mismatch phases remain the same, as well as the one-day structure. The critical change to this version is the addition of item evaluation to the test phase; both the recall of sequence order and specific image identification are tested. Subjects, again, are cued with the first item in a sequence and are now presented with three stimulus-quartet boxes. Each stimulus-quartet box contains a tiled grid of four images of different, but strongly visually similar, exemplars of the items that made up the rest of the sequence, including the actual image seen previously, and three lures. Subjects are instructed to first rebuild the remaining sequence using the stimulus-quartet box. Once they have confirmed their sequence order, subjects are prompted to select the image within each stimulus-quartet box they encountered previously. As in the previous version a total of 50 mismatch, 40

repeat, and 30 new sequences are presented and tested; 360 individual items were tested, 3 items (positions 2-4) per each of 120 sequences. This version did not present any lures for the sequence reconstruction, as the incorrectly rejected items would not be tested, and/or lures incorrectly included in rebuilt sequences could be tested twice. Figure 4.3 shows the design for this experiment.

Participants

Seventeen (11 female) subjects participated in the second Experiment ($mean_{age}$ = 22.76 years). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time.

Results

The accuracies of sequence reconstruction were very similar to Experiment 1: 47.99% correct overall (*SD* = 15.47%), 75.44% correct for repeats (*SD* = 26.61%), 40.20% correct for novels (*SD* = 17.30%), 30.71% correct (*SD* = 10.44), and 49.29% in the old order for mismatches (*SD* = 16.15%). The removal of lure items in this design explains the overall increase in accuracies as compared to Experiment 1, and also raises the chance rate to 16.66%.

The addition of the item test did not result in any evidence of increased item recall for those presented in mismatch sequences. Overall, items seen in repeat (*mean* = 77.30%, *SD* = 19.56%) did not differ from those seem in mismatch sequences (*mean* = 77.06%, *SD* = 18.83%) ($t_{(16)}$ = 0.16, p = 0.87). This held true when looking at all positions of items: those items presented in the second position, before any mismatch, were recalled similar between repeat (*mean* = 76.47%, *SD* = 21.40%) and mismatch

(*mean* = 75.76%, *SD* = 22.33%) ($t_{(16)}$ = 0.34, p = 0.74). Items seen in the third and fourth position during the mismatch phase were the same for repeat (3rd: *mean* = 79.56%, *SD* = 18.78%; 4th: *mean* = 75.88%, *SD* = 19.80%) and mismatches (3rd: *mean* = 77.88%, *SD* = 17.66%; 4th: *mean* = 77.53%, *SD* = 17.37%) (3rd: $t_{(16)}$ = 1.02, p = 0.32; 4th: $t_{(16)}$ = 0.73, p = 0.47). See Figure 4.2.

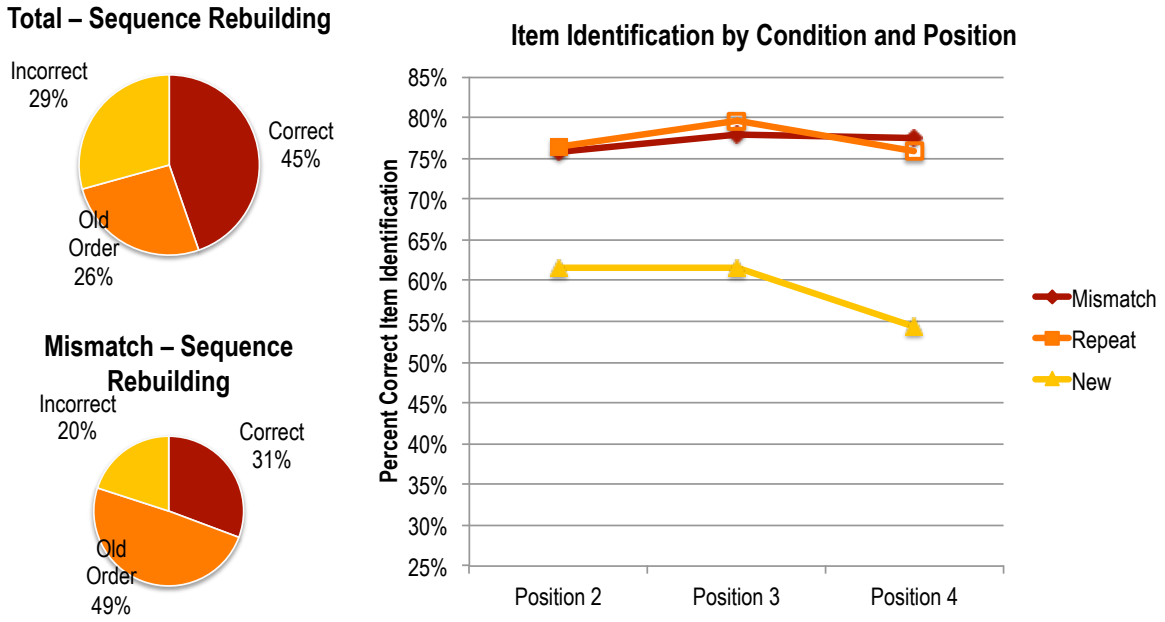
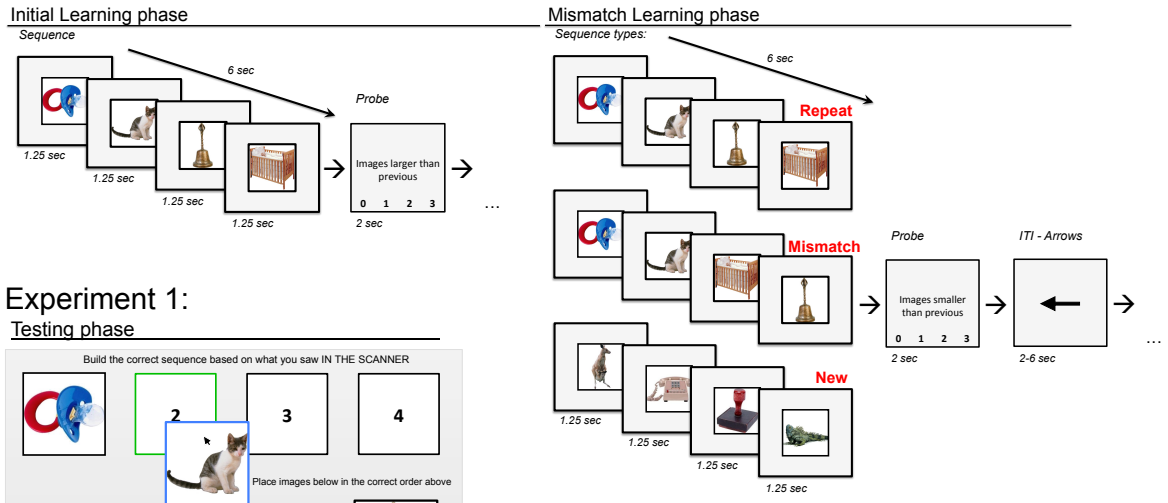


Figure 4.2: Results for Experiment 2. Both sequence rebuilding and items were tested. Open squares indicate items presented out of order.

Discussion

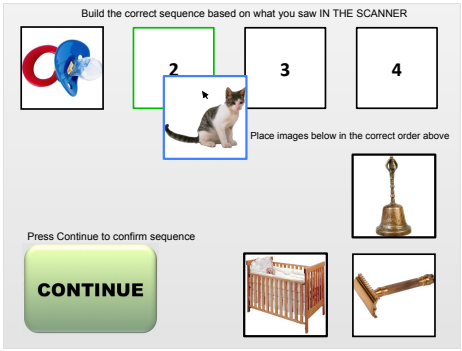
Results here did not show our expected benefit for items whose presentation order had been swapped, nor any differences between our two conditions of interest. Given the overall high level of recall for all items, upward of 75% in all conditions and positions, it is possible that our task provided too much context during test, which boosted memory in all conditions. In Experiment 3 we removed each item from its sequence-context, and move the test to the next day in order to reduce overall recall. These changes will hopefully allow for differences between conditions to emerge.

Learning phases, Experiment 1-3:



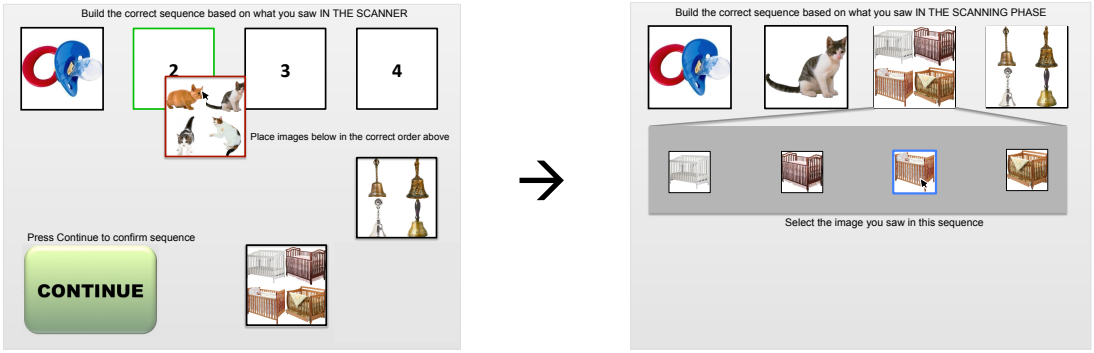
Experiment 1:

Testing phase



Experiment 2:

Testing phase



Experiment 3-8:

Testing phase

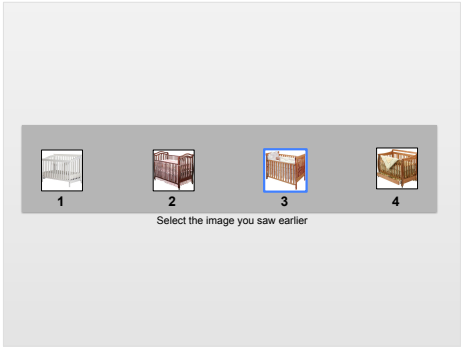


Figure 4.3: Experimental Design, part 1. Learning phases for experiments 1-3, and test phases for all experiments

Experiment 3: Item Only Test

Design

This Experiment eliminated the sequence order rebuilding from the test phase; the learning and mismatch phases remain the same. During the test phase subjects are presented with four images of different exemplars of items that were previously presented in the 2nd, 3rd, or 4th position of the sequences (the 1st item of the sequences are not include in the testing set), including the actual image seen previously, and three lures. The exemplar set is presented individually from the rest of its sequence set, and the order of items tested was randomly selected from the entire testing set. Again, 360 individual items were tested, and condition counts remained the same as the previous in experiments. Along with the change in the test phase, this experiment divided the task into two days, with the test phase occurring a day after the learning and mismatch phases. Figure 4.3 shows the design for this experiment.

Participants

Twenty-six (20 female) subjects participated in the second Experiment ($mean_{age}$ = 19.77). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time. One subject was excluded due to experimenter error, and two subjects did not return for the second day of the experiment, so this dataset included 23 (17 female) participants.

Results

Removing the item test from its sequence-context, and allowing a daylong delay before the phase test did result in differences between images seen in repeat and

mismatch sequence conditions. Results here showed greater recall for items in repeat sequences than those in mismatch sequences. Overall the repeat condition, regardless of position, (*mean* = 62.86%, *SD* = 19.97%) were recalled more than mismatches (*mean* = 59.54%, *SD* = 19.43%) ($t_{(22)}$ = 3.60, p = 0.002).

When we break the conditions down based on position we find, again, that items seen in repeat sequences are recalled at a greater rate than those images seen in the mismatching sequences, although these differences were not significant in all positions. Items presented in the second position, before any mismatches, repeats (*mean* = 64.02%, *SD* = 20.38%) were remembered more than mismatches (*mean* = 60.17%, *SD* = 23.19%) ($t_{(22)}$ = 2.13, p = 0.044). The benefit for repeat sequences' items (*mean* = 63.37%, *SD* = 18.81%) over mismatched (*mean* = 58.61%, *SD* = 20.26%) is significant when looking at the third position ($t_{(22)}$ = 3.00, p = 0.007). Finally, recall for items seen in the fourth position of sequences did not differ between repeat (*mean* = 61.20%, *SD* = 23.19%) and mismatch (*mean* = 59.83%, *SD* = 20.26%) ($t_{(22)}$ = 0.57, p = 0.57) conditions.

It is worth noting that the difference in recall for items presented in the third position of repeat verse mismatch sequences is possibly an artifact of the apparent linear decrease in item recall by position. Within repeat sequences, items in the 2nd position (*mean* = 64.02%, *SD* = 20.38%) are recalled more, but not significantly more than items in the 4th position (*mean* = 61.20%, *SD* = 23.19%), $t_{(22)}$ = 1.14 p = 0.27. When we compare recall for items based on the position they were seen during the three learning exposures, the decrease in recall for items in the 3rd position of mismatch sequences (*mean* = 59.83%, *SD* = 20.26%) compared to repeat sequences (*mean* =

63.37%, *SD* = 18.81%) is much reduced, and no longer significant ($t_{(22)}$ = 1.79, p = 0.086). The difference between recall for items of repeat (*mean* = 61.20%, *SD* = 23.19%) and mismatch (*mean* = 58.61%, *SD* = 20.26%) sequences in the fourth position at learning does increase, relative to this comparison using the mismatch phase's order, but the difference is still not significant ($t_{(22)}$ = 1.51, p = 0.144). See Figure 4.4.



Figure 4.4: Results from Experiment 3. Only images were tested. Open squares indicate items presented out of order. Data presented both in original presentation order, and order as presented during the mismatch phase.

Discussion

These results indicate that consistency, rather than mismatches, leads to greater item recall in our task design. Rather than benefiting the memory for items that mismatch their original order, the violation seems to disrupt some of the encoding of the items in that sequences, particularly it seems, for items that precede the mismatch. The mechanism of this disruption is not clear from our experiment, and could result from a number of factors. The swapping of items within a sequence is likely fairly attention grabbing, and this attentional shift, rather than benefiting its source seems to result in a cost to the material that are attended away from. Given that the greatest difference in

item recall between repeat and mismatch sequences is in the second position (when examining sequences based on the order they were seen in the majority of presentations), right before the presentation of mismatches, it would indicate that the encoding of these items is disrupted by the mismatches that follow. Alternatively, or in combination with, the confirmation of predicted material could result in a stronger memory benefit, at least in our well-learned sequences, than a violation.

Whatever the mechanism behind our results here, it is clear that the presentation of mnemonic mismatches in the form of swapped sequence positions does not result in a benefit to item encoding. Previous research has shown that such a design does, in fact, cause increased activity within the CA1 subfield of the hippocampus in response to the order violation (Chen et al., 2015; Kumaran & Maguire, 2006), but our results indicate that this activity is not sufficient to boost encoding over and above the benefit of multiple consistent presentations of material. However, this result does not indicate that there are no benefits to be found to encoding resulting from mnemonic mismatches, rather that the comparison of mismatching items with well-learned material is not ideal for the investigation of this effect. As by the time the mismatch occurs all images had been seen three times previously it is likely that the majority of item encoding was already completed. Any mismatch benefit therefore would only apply to the last learning event, and it seems to not be a sufficiently large effect to overcome the benefit of consistency. This limitation would have applied to the behavioral study design used by Chen et al. (2015) to evaluate the behavioral results of mismatches, and can possibly explain their relatively weak result.

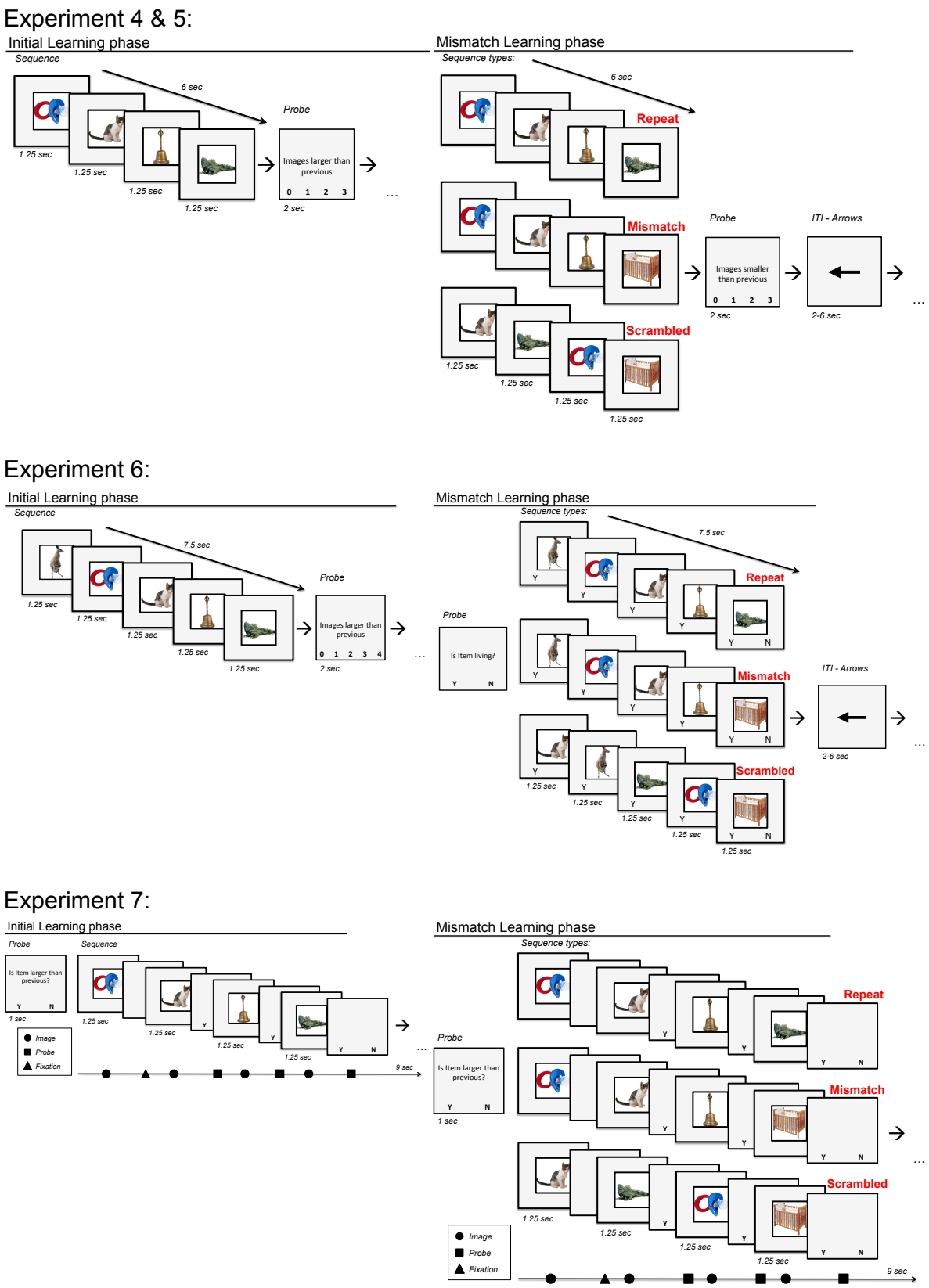


Figure 4.5: Experimental Design, part 2. Learning phases for experiments 4-7

To limit all item learning of the critical violation items to the mismatch event alone, as well as increase the degree of mismatch, in the next experiment we changed what comprises a mismatch. Rather than switching the order of items at the end of a sequence we will present a novel item in place of one of the items previously presented. Also, rather than comparing a single instance of a mismatch to items seen four times in a consistent manner, we will compare novel items in otherwise intact sequences, where the novel item will violate the expectation generated by the presentation of the rest of the sequence, to novel items presented in scrambled sequences, where no such expectation would be triggered.

Experiment 4: Novel Items as Mismatch

Design

Experiment 4 represents the largest change in task design from previous versions. The task keeps the three phase format, and the test phase is conducted a day following the other two phases. This test phase again only tests item (and not sequence) memory. Critically, what constitutes a mismatch has changed in this version. Rather than swapping the position of items, the mismatch trials present a novel, never seen before image, in the 3rd or 4th position in the sequence, while maintaining the learned order of the remaining items. Furthermore, instead of the presentation of "new" sequences, "scrambled" sequences are presented. These sequences will have been seen in the initial learning phase, but presented out of order during the mismatch phase such that no item occurs in its original position nor does any item follow the item that originally preceded it. Like the mismatch sequences, these scrambled trials present a

novel image in the 3rd or 4th position. It is these scrambled novel images to which the mismatch novels are compared. The scrambling of the sequence provides a control for familiarity of the items presented with the novel item, without inducing a strong prediction of the item that should be presented next, as should be triggered in our mismatch condition. By using a novel image to violate the expectations triggered by the early items of a sequence, rather than a swapping of items, will allow us to attribute any benefit for the recall of these items to its condition during the mismatch phase, as this is the only time the image was seen prior to testing.

To maintain the expectation of complete sequences repeat trials are still included in the mismatch phase. In this Experiment design 35 mismatch, 35 scrambled, and 25 repeat sequences are presented for a total of 95 sequences, and 285 individual items were tested. Finally, the n-back task was removed from the mismatch phase, as no images are pre-exposed. Figure 4.5 shows the design for this experiment.

Participants

Thirty-four subjects (22 female) participated in this Experiment ($mean_{age}$ = 21.00). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time. Four subjects were excluded due to performance below chance on overall recall on novel items, and one subject did not return for the second day test, so this dataset includes 29 participants (18 females).

Results

Consistent with Experiment 3, non-novel items seen in intact repeated sequence (*mean* = 71.68%, *SD* = 15.81%) tended, although not significantly, to be recalled more

than non-novel items seen in sequence that presented a novel item (mismatch) (*mean* = 69.14%, *SD* = 17.48%) ($t_{(28)}$ = 1.93, p = 0.063). Interestingly, the recall for non-novel items in repeat sequences did not differ from the recall rates of non-novel items in a scrambled sequence (*mean* = 70.94%, *SD* = 16.61%) ($t_{(28)}$ = 0.58, p = 0.561). Likewise, the recall rates for non-novel items did not differ between the conditions that presented novel items, mismatch and scrambled ($t_{(28)}$ = 1.19, p = 0.242).

The omnibus test of successful recall for novel images between conditions, intact (mismatch) verse scrambled order, and novel image position within a sequence, either 3rd or 4th position, resulted in a significant Condition * Position interaction ($F_{(1,28)}$ = 9.071, *MSE* = 0.017, p = 0.005). However, this test revealed no main effect of position ($F_{(1,28)}$ = 0.685, *MSE* = 0.013, p = 0.415), nor condition ($F_{(1,28)}$ = 2.11, *MSE* = 0.010, p = 0.157), indicating that overall novel items seen in a mismatch sequence (*mean* = 46.50%, *SD* = 14.98%) were not significantly recalled more than those seen in a scrambled sequence (*mean* = 43.97%, *SD* = 14.19%).

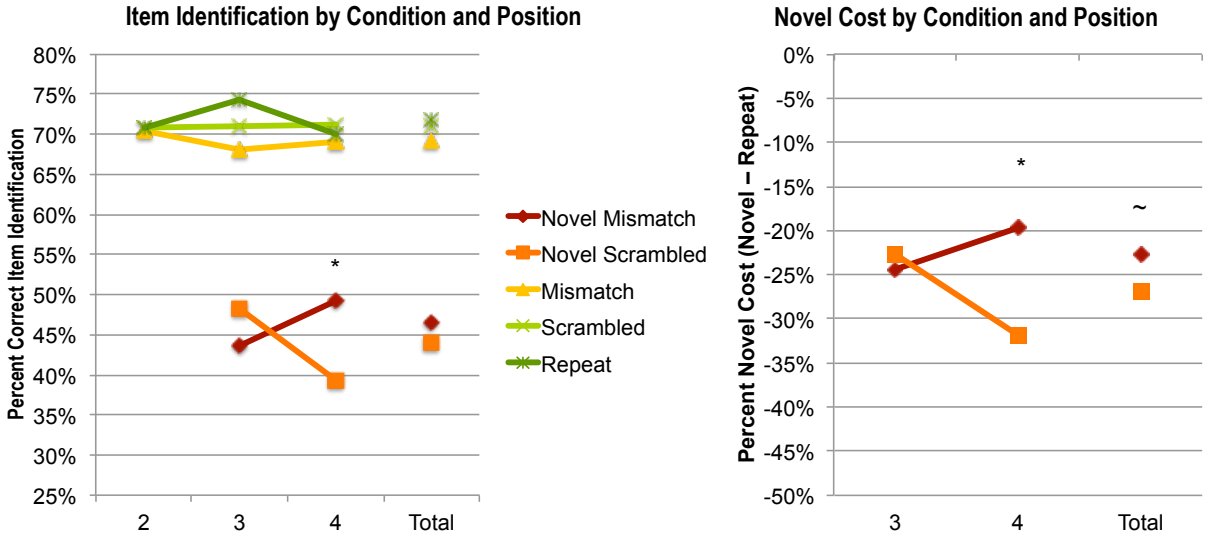


Figure 4.6: Results from Experiment 4. Novel items presented as the mismatch. Both recall rates, and novel costs shown.

The interaction results from increased recall of novel items seen in the last, 4th, position of mismatch sequences (*mean* = 49.23%, *SD* = 16.58%) relative to the same position in scrambled sequences (*mean* = 39.24%, *SD* = 16.37%) ($t_{(28)}$ = 3.01, p = 0.005), while no difference was observed between these two conditions for novel items presented in the penultimate, 3rd, position (mismatch: *mean* = 43.61%, *SD* = 17.26%; scrambled: *mean* = 48.28%, *SD* = 18.31%, $t_{(28)}$ = 1.71, p = 0.098). Furthermore, in mismatch sequences novel items presented in the 4th position (*mean* = 49.23%, *SD* = 16.58%) were recalled more, although not quite significantly, than items presented in the 3rd position (*mean* = 43.61%, *SD* = 17.26%) ($t_{(28)}$ = 1.92, p = 0.064), indicating that the presentation of more items of a sequence before a mismatch event increases the mismatch benefit. Conversely, in scrambled sequences novel items at the end of a sequence (*mean* = 39.24%, *SD* = 16.37%) are recalled less than novels seen in the 3rd position (*mean* = 48.28%, *SD* = 18.31%) ($t_{(28)}$ = 2.60, p = 0.014). While novel items were not overall recalled more in mismatch sequences, when these items were seen last there was a significant recall benefit resulting from the violation of the otherwise intact sequence.

To control for slight variations in overall recall for items between position and condition we converted our novel recall rates into novel costs. We subtracted each subjects' recall rate for non-novel items in each position for each condition from the corresponding novel recall rates. This resulted in an estimate of the cost to later recall of an item being novel relative to the recall of old items in that condition. The novel-cost of mismatch sequences (*mean* = -22.05%, *SE* = 2.51%) is less, although just failing to pass significance, than the novel-cost of scrambled sequences (*mean* = -27.35%, *SE* =

3.14%) ($F_{(1,28)}$ = 3.959, *MSE* = 0.021, p = 0.056). While the overall recall rate benefit for novel items in mismatch sequences was not significantly greater than in scrambled sequences, being novel was less detrimental for the mismatch condition relative to the scrambled. The Condition * Position interaction was still significant for novel costs ($F_{(1,28)}$ = 5.110, *MSE* = 0.028, p = 0.032), and there was no main effect for position ($F_{(1,28)}$ = 0.370, *MSE* = 0.041, p = 0.548). Again, this interaction results from decreased novel cost for items seen in the last position of mismatch sequences (*mean* = -19.73%, *SD* = 16.22%) relative to the same position in scrambled sequences (*mean* = -31.86%, *SD* = 21.90%) ($t_{(28)}$ = 2.62, p = 0.014), while no difference was observed between the novel cost for novel items presented in the 3rd position (mismatch: *mean* = -21.40%, *SD* = 19.35%; scrambled: *mean* = -22.62%, *SD* = 22.94%; $t_{(28)}$ = 0.51, p = 0.508). See Figure 4.6.

Discussion

We show here a strong benefit for the subsequent recall of items that mismatched previously encounters, as compared to items that did not violate a mnemonic prediction. Novel images that replaced, and thus violated the expectation of what was expected based on previous experience, the last item of an otherwise intact sequence were recalled at a higher rate than novel images placed in the same position of a sequences where all old items where presented out of order, thus reducing or eliminating the mnemonic prediction. The cost to recall of novel items relative to condition matched old items was less for the mismatch than scrambled conditions. These results indicate that reduction in recall rates for items seen only once compared to items seen multiple times in the same condition is mediated when that single viewing

instance represents a mnemonic violation. These results support the body of research showing sensitivity within the hippocampus, and the CA1 in particular, to mismatches (Chen et al., 2015, 2011; Duncan et al., 2012; Kumaran & Maguire, 2006, 2007c), and the theory that places this hippocampal response as a trigger to encode information in long-term memory (Lisman & Grace, 2005).

A number of factors may contribute to our result that novel recall rates only differed between condition in the 4th and final position of sequences. Due to the design of the task items seen in last are only evaluated once, a relationship judgment is made about this item relative to the item that preceded it, while items seen in the second and third positions are evaluated twice, both to the item that preceded it, and again when the next item is evaluated. This difference in number of evaluations potentially explains the decrease in novel recall rates within the scrambled condition in the last position. There was no such decrease for novel items in this last position of mismatch sequences; rather there was a trend towards an increased recall rate from the 3rd position. The benefit from violating mnemonic expectations appears to overcome the reduction seen in scrambled sequences due to reduced mental engagement with the last items, although the lack of a mismatch benefit in the 3rd position might indicate that this benefit is fairly small and the two mental engagements in this position are enough to overwhelm it.

An additional interpretation of the mismatch benefit in the final position only is that the benefit is greatest when prediction strengths are greatest: after the presentation of three items of a sequence. While previous fMRI studies have found CA_1 responses to mismatches after the presentation of two items of a four item sequence (Chen et al.,

2015; Kumaran & Maguire, 2006) it is possible that those responses were fairly insignificant compared to other factors influencing encoding, and the presentation of more of the intact sequence would have increased the CA_1 response. Experiment 5 will seek to evaluate the effects of memory strength has on the mismatch benefit, which has been shown to influence the response to mnemonic violation in the CA_1 (Chen et al., 2015). We will follow up on the idea that when a mismatch occurs later in a sequence it results in a greater recall benefit for novel items in Experiment 6, where we present participants with five item sequences.

Experiment 5: Effect of Memory Strength

The first variant to the design used in Experiment 4 will test effects of over all strength of memory on mismatch recall benefits. Chen et al. (2015) showed increased CA1 response to mismatches when the mismatch occurred in a sequence presented several times previously when compared to sequences presented fewer times. To tie our behavioral results more concretely to the activity within the CA_1 we will here replicate Chen et al.'s (2015) manipulation of pre-mismatch exposures to see if violation of better learned sequences results in stronger encoding of novel images than less well learned ones.

Design

Experiment 5 follows closely the design of Experiment 4, including the day delay between learning and testing phases, the use of novel items replacing items in sequences to trigger mnemonic mismatches, and the use of scrambled sequences as the comparison condition. To vary the strength of the memories for sequences, and thus vary prediction strength, here half of sequences, in all three conditions, are

presented four times prior to the mismatch phase (Strong sequences), and the other half of sequences are presented only once (Weak sequences). The sequences seen a single time were evenly distributed between the four learning blocks, so in each learning block 75% of sequences had been seen previously (or would be seen again for the first block), while 25% were novel sequences, not to be seen again till the mismatch phase. The additional learning phase, relative to previous studies, asked participants to determine how many of the items in that series were "more valuable than" the item that preceded it. As we saw the greatest difference between the mismatch and scrambled condition in the fourth position of sequences we only present novel items in this last position here in order to maintain the same power between exposure numbers as between positions in Experiment 4. All sequences are presented during the mismatch phase, and the 2nd, 3rd, and 4th items from all sequences as seen during this phase were tested the next day. The number and distribution of sequences per condition is the same as in Experiment 4. Figure 4.5 shows the design for this experiment.

Participants

Twenty-six subjects (17 female) participated in this Experiment ($mean_{age}$ = 21.00 years). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time.

Results

Unlike Experiments 3 and 4 the overall recall for old items seen in mismatch sequences (*mean* = 69.67%, *SD* = 14.78%) did not differ from items seen in repeat sequences (*mean* = 69.38%, *SD* = 14.92%) ($t_{(25)}$ = 0.204, p = 0.840). Although, items seen in scrambled sequences (*mean* = 66.04%, *SD* = 14.73%) were recalled less than

either repeat ($t_{(25)}$ = 2.445, p = 0.022) or mismatch ($t_{(25)}$ = 2.211, p = 0.036) sequences. When recall rates were further broken down by number of previous presentations, the difference between scrambled and the other two conditions was only seen when sequences were strongly learned. Items presented in scrambled sequences seen only once before the mismatch phase (*mean* = 58.01%, *SD* = 14.64%) were not recalled less than items seen in repeat sequences (*mean* = 61.11%, *SD* = 15.33%) ($t_{(25)}$ = 1.57, p = 0.128) nor mismatch sequences (*mean* = 61.32%, *SD* = 16.06%) ($t_{(25)}$ = 1.470, p = 0.154). However, items presented in scrambled sequences seen four times prior to the mismatch phase (*mean* = 74.54%, *SD* = 16.04%) were recalled less that items seen the same number of times in mismatch (*mean* = 78.51%, *SD* = 119.92%) ($t_{(25)}$ = 2.11, p = 0.045) and trended to be recalled less than repeat sequences (*mean* = 77.02%, *SD* = 15.59%) ($t_{(25)}$ = 1.75, p = 0.092).

In contrast with our hypothesis that the benefit in subsequent recall for novel items presented in mismatch sequences would increase as the strength of the sequences increased, the omnibus test of successful recall for novel images between conditions, mismatch verse scrambled, and the strength of memory of the sequences, weak verse strong, did not result in any differences. Recall rates did not vary between mismatch (*mean* = 44.67%, *SE* = 2.35%) and scrambled (*mean* = 41.81%, *SE* = 2.30%) ($F_{(1,25)}$ = 1.036, *MSE* = 206.92, p = 0.318), nor between weak (*mean* = 43.26%, *SE* = 2.26%) and strong (*mean* = 43.21%, *SE* = 2.29%) sequence strength ($F_{(1,25)}$ = 0.083, *MSE* = 180.39, p = 0.983). Furthermore, the interaction between condition and presentations was not significant ($F_{(1,25)}$ = 5.110, *MSE* = 1.652, p = 0.210). Novel images seen in mismatch sequences were not recalled more when seen in strong

sequences (*mean* = 42.99%, *SD* = 16.54%) than when seen in weak sequences (*mean* = 46.36%, *SD* = 14.14%) ($t_{(25)}$ = 0.89, p = 0.379), and the same is true for novel items seen in scrambled sequences (Weak: *mean*= 40.17%, *SD* = 15.02%; Strong: *mean* = 43.43%, *SD* = 14.79%) ($t_{(25)}$ = 0.91, p = 0.373). Furthermore, there was no difference between recall for novel images seen in mismatch (*mean* = 42.99%, *SD* = 16.54%) or scrambled (*mean* = 43.43%, *SD* = 14.79%) sequences seen four times previously ($t_{(25)}$ = 0.11, p = 0.913), although there was a slight, but not significant, benefit for novel images seen in mismatch sequences (*mean* = 46.36%, *SD* = 14.14%) compared to when seen in scrambled sequences (*mean*= 40.17%, *SD* = 15.02%) when those sequences had previously been presented only once ($t_{(25)}$ = 1.76, p = 0.090).

When recall rates were converted to novel costs (calculated by subtracting recall rates for non-novel item from the novel recall rates; see Experiment 4 results for a more detailed description) there was, unsurprisingly, a large increase in novel cost for strong (*mean* = -33.31%, *SE* = 2.31%) verse weak (*mean* = -16.40%, *SE* = 2.49%) sequences ($F_{(1,25)}$ = 42.09, *MSE* = 179.73, p < 0.001), reflecting the large increase in recall rates for

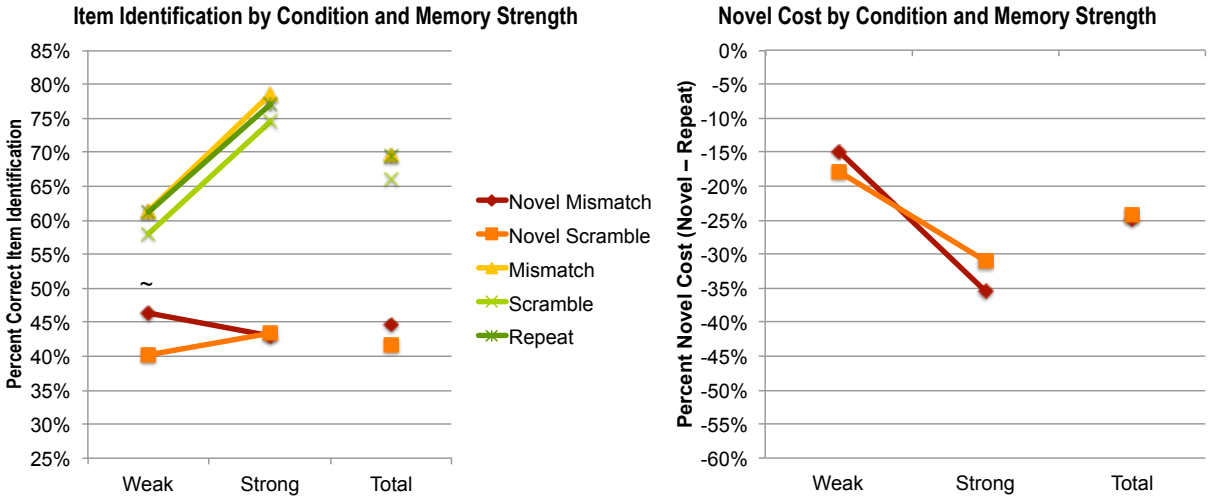


Figure 4.7: Results for Experiment 5. Novel items presented as the mismatch, sequences presented either once (weak) or four times (strong) prior to mismatch presentation. Both recall rates, and novel costs shown.

non-novel item with more presentations. Other than this expected effect there were no differences due to condition (mismatch: *mean*, = -25.24%, *SE* = 2.54%; scramble: *mean* = -24.48%, *SE* = 2.31%; $F_{(1,25)}$ = 0.079, *MSE* = 190.93, *p* = 0.780), nor did we find a Condition * Strength interaction ($F_{(1,25)}$ = 1.789, *MSE* = 193.38, *p*=0.193). Novel costs did not differ between mismatch (weak: *mean* = -14.96%, *SD* = 16.31%; strong: *mean* = -35.52%, *SD* = 16.08%) and scrambled (Weak: *mean* = -17.84%, *SD* = 15.30%; Strong: *mean* = -31.11%, *SD* = 15.00%) conditions when sequences had been seen once ($t_{(25)}$ = 0.78, *p* = 0.442) or four times ($t_{(25)}$ = 1.11, *p* = 0.279) previously. See Figure 4.7.

Discussion

Here we failed to find our expected result that the mismatch benefit seen in Experiment 4 would be greater for novel images replacing the last item of a well learned sequence then novel items in a weakly learned sequence. The results of this experiment only demonstrated a weak trend towards an increased novel recall rate for mismatch sequences over scrambled when those sequences had only been presented once prior. There was no difference between these two conditions for the sequences seen several times, nor were there any differences between the recall rates of novel items in mismatch condition based on sequence strength.

These results are surprising given the findings of Chen et al. (2015), which showed an increased response within the CA_1 subfield of the hippocampus to violations in sequences with stronger memory strength compared to weaker ones. Not only did we not show an increase of mismatch benefit for more well learned sequences, the overall recall rate for these novel images seen in the mismatch condition was lower in the strong condition, although not significantly, than in the weak. Our results would

indicate that stronger mismatch signaling in the hippocampus does not lead necessarily to greater encoding. In Lisman and Grace (2005), the authors emphasize the complexity of the system that modulates the release of dopamine back into the hippocampus after the prediction violation had been detected. This downward arc of the Hippocampal-VTA loop includes nodes in the accumbens, pallidum, and prefrontal cortex, which convey goal and salience information, information that helps determine what is important to encode into long term memory. Given the overall lack of significant results in this experiment it is difficult to interpret what factors might have contributed to our findings, but it is possible that while mismatches in well learned materials lead to greater hippocampal responses, there were greater motivational elements for encoding during the presentation of the weak sequences. The second presentation of a sequence might have been inherently more interesting or automatically triggered more overall learning than ones seen multiple times.

To control for these possible unknown motivational/salience differences between the viewing of sequences with different number of previous presentations, as well as to follow up with our position effects seen in Experiment 4, we next conducted an experiment where prediction strength was modulated by the position within a sequence.

Experiment 6: Effect of Prediction Strength

As the results of Experiment 4 show that the position in a sequence a novel image was presented affected the observed mismatch benefit we here further evaluate this result by lengthening sequences to five items, and present the novel image at the end of this longer sequence to compare to the mismatch benefit for novel items presented earlier. The mismatch signal in the CA_1 follows from a violation of a prediction

generated upstream in the CA_3, so the presentation of more items in a learned sequences should generate a stronger CA_3 prediction, and hence a greater mismatch response, and thus greater encoding. This manipulation of prediction strength keeps the number of previous presentations constant, which eliminates the possible confounds that might have lead to our unexpected results in Experiment 5. We expect that when a novel image is presented in the fifth and final position of an otherwise intact sequence that novel will be recalled at a greater rate than either novels presented in earlier positions or in a scrambled sequence.

Design

This Experiment presents the learning phase as in Experiment 4, with three exposures of all sequences, however, rather than presenting four items, sequences are five items long. During the learning phase participants still judge how many of the items in that series met a specified relationship with the item that preceded it, although responses can now range from '0' times a relationship was met, to '4' times. An additional change is that during the mismatch phase participants make semantic judgments about each item individually, replying 'yes' or 'no' to the question "larger than a shoebox." This non-relationship judgment was employed in order to keep the number of evaluations of each item consistent between positions. In previous versions items seen in the last are only evaluated once, a relationship judgment is made about this item relative to the item that preceded it, while items seen in the second and third positions are evaluated twice, both to the item that preceded it, and again when the next item is evaluated. This difference in number of evaluations could have cause the decrease in recall rates for novel images seen in the last position of scrambled

sequences revealed in Experiment 4. Keeping the number of evaluations constant is necessary to attribute any differences in subsequent recall between positions to different prediction strengths. The number and distribution of sequences per condition was the same as in Experiment 4, while the number of individual items tested increased to 380, reflecting the additional item per each of the 95 sequences. Figure 4.5 shows the design for this experiment.

Participants

Nineteen (8 Female) participants took part in this Experiment ($mean_{age}$= 20.53 years). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time. One subject did not return for day 2 of the Experiment, and a computer error prevented collect of data from a second subject, so a total of 17 subjects (7 Female) make up this data set.

Results

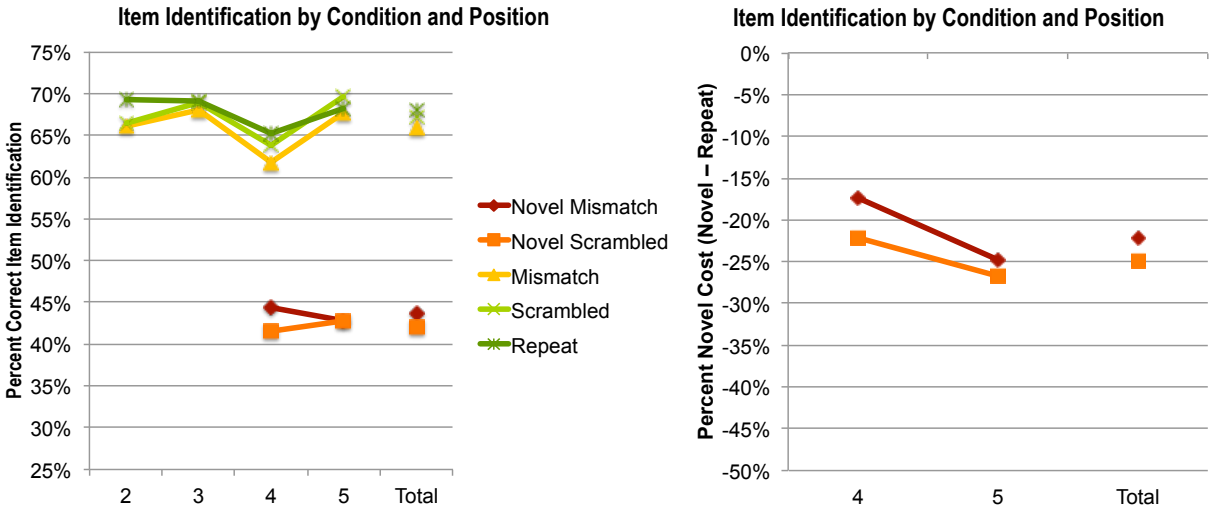


Figure 4.8: Results for Experiment 6. Novel items presented as the mismatch, 5 item sequences Both recall rates, and novel costs shown

The results of this experiment did not support our hypothesis that the benefit to the recall of novel items seen after more of a learned sequence would be greater than when less of a sequence had been presented. Recall for these novel items did not differ based on condition (Mismatch: *mean* = 43.62%, *SE* = 3.05%; Scrambled: *mean* = 42.21%, *SE* = 3.81%) ($F_{(1,16)}$ = 0.305, *MSE* = 0.011, *p* = 0.588), nor position ($F_{(1,16)}$ = 0.004, *MSE* = 0.013, *p* = 0.949), and there was no Condition * Position interaction ($F_{(1,16)}$ = 0.405, *MSE* = 0.008, *p* = 0.533). Similarly, the novel cost (calculated by subtracting recall rates for non-novel item from the novel recall rates. See Experiment 4 results for a more detailed description) did not differ based on condition (Mismatch: *mean* = -21.03%, *SE* = 2.82%; Scrambled: *mean* = -24.56%, *SE* = 3.38%) ($F_{(1,16)}$ = 0.972, *MSE* = 0.022, *p* = 0.339), nor position ($F_{(1,16)}$ = 2.267, *MSE* = 0.028, *p* = 0.152), and there was no Condition * Position interaction ($F_{(1,11)}$ = 0.147, *MSE* = 0.023, *p* = 0.706). Overall, there was no difference between the recall rates for non-novel items presented in Repeat (*mean* = 67.97%, *SD* = 12.13%) and Mismatch (*mean* = 65.85%, *SD* = 13.53%) sequences ($t_{(16)}$ = 1.25, *p* = 0.229), nor between Repeat and Scrambled (*mean* = 67.20%, *SD* = 13.22%) sequences ($t_{(16)}$ = 0.375, *p* = 0.712), or between Mismatch and Scrambled sequences ($t_{(16)}$ = 0.714, *p* = 0.486). See Figure 4.8.

Discussion

Here, we again fail to show that greater prediction strength results in stronger encoding of items that violate these predictions. No differences were detected between novel recall rates, nor novel costs, based on condition or position. As in Experiment 6, the lack of significant results makes interpretations difficult. One possibility is that in order to control for the amount of mental engagement with each item in the sequence

we changed the task during the mismatch phase to a simple semantic judgment rather than an integrative relationship judgment. This semantic judgment might have reduced the triggering of predictions as this task deemphasized the overall sequence. Additionally, the requirement of a response to each item might have distracted participants from any predictions that were being generated. The semantic judgments might also have itself been novel enough compared to the previous task to give a mismatch benefit across all conditions, resulting in the observed lack of differences for novel recall rates between conditions. Whatever the cause of our results here, and those of the previous experiment, it is clear that increasing prediction strength, either through the manipulation of the number of presentations of a sequences or the position within a sequence that a violation occurs, does not directly lead to increased encoding. These two experiments would indicate that other factors such as task demands play a larger roll in what is learned than prediction strength.

Experiment 7: Response Prediction Mismatch

To eliminate any task novelty from our design this experiment used the same task through both learning and mismatch phases, only asking one integrative relational judgment multiple times. This experiment adds the element of predicted response violation to the visual mismatch, so as to see if this task oriented mismatch boosts the mismatch benefit.

Design

All versions of the task up to this point have asked participants to make several different integrating relationship judgments during the learning phase. Subjects kept a running total the number of times each new relationship was met within a sequence,

and replied at the end of each sequence. In this experiment participants now respond after each image in a series, yes or no to the question: “Is this item larger than the item that came before it.” This one probe question was used in all three learning blocks as well as during the mismatch phase. Rather than a fixation cross between each item in the four-item sequence, a “Yes/No” probe was presented after the second, third, and fourth items (as there is no comparison to make after the first item a fixation cross was presented after these items). Each probe is on the screen for 1 second before moving on to the next image, so as to allow participants enough time to respond to the probe while still maintaining a pace that would build a representation of the images as a sequence and lead to a strong prediction of the next image. Participants reply ‘1’ to indicate ‘yes’ that the current item was larger than the item that came before it, and ‘2’ to indicate ‘no’ to this probe. As in Experiment 5, we only present novel images in the fourth and final position as this was were we found the largest differences due to condition in Experiment 4. Other than these changes, this task resembles that of Experiment 4, with four-item sequences, and the same number and distribution of sequences per condition, and number of items tested on the second day. Figure 4.5 shows the design for this experiment.

Participants

Seventeen (12 Female) participants took part in this Experiment (mean age= 21.14 years). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time. A computer error prevented collection of data from one subject, so a total of 16 subjects (11 Female) make up this data set.

Results

Results here failed to show a replication of the benefit in recall for novel items presented in an otherwise intact sequence over novel items seen in sequences presented out of order. Novel items presented in mismatch sequences (*mean* = 42.88%, *SD* = 13.57%) were not recalled more than items seen in scrambled sequences (*mean* = 46.01%, *SD* = 16.23%) ($t_{(15)}$ = 0.88, *p* = 0.392). When non-novel recall rates were subtracted from novel recall rates, resulting in a novel cost, this cost also did not differ between mismatch (*mean* = -25.43%, *SD* = 11.45%) and scrambled (*mean* = -24.48%, *SD* = 12.74%) conditions ($t_{(15)}$ = 0.26, *p* = 0.800). Furthermore, as in Experiment 6, no difference emerged between recall rates for non-novel images based on condition. Non-novel recall rates were not greater for repeat sequences (*mean* = 71.92%, *SD* = 15.69%) relative to mismatch sequences (*mean* = 68.32, *SD* = 15.05%) ($t_{(15)}$ = 1.29, *p* = 0.186), nor scrambled sequences (*mean* = 70.49%, *SD* = 15.64%) ($t_{(15)}$ = 0.88, *p* = 0.391). Likewise, non-novel recall did not differ between mismatch and scrambled sequences ($t_{(15)}$ = 1.09, *p* = 0.292). See Figure 4.9.

Discussion

This Experiment fails to show an increased mismatch benefit when mismatches also violated the expected response to that image. Although this design controlled for the change

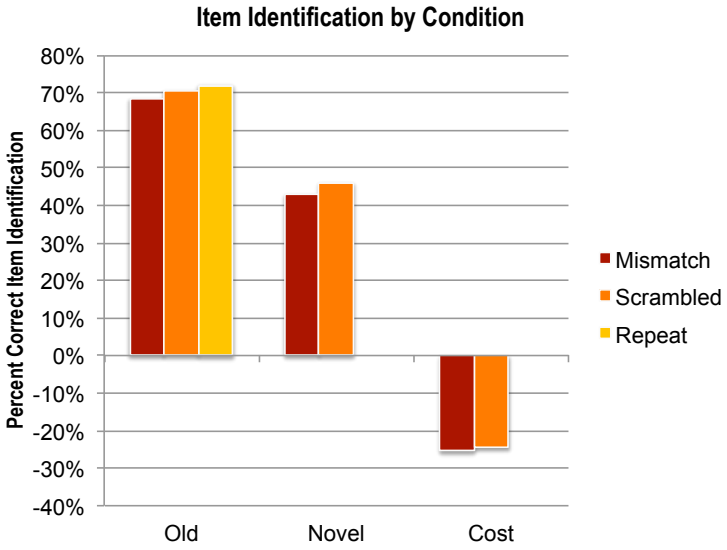


Figure 4.9: Results for Experiment 7. Novel items presented as the mismatch, only one relationship probe presented, judgments made after ever image. Both recall rates, and novel costs shown

in tasks between phases as a confounding source of novelty, seen in the previous experiment, there were several elements of our design here that might have undermined our ability to find a mismatch benefit. Foremost, the scrambled condition, which acted as the comparison condition in previous versions, was no longer an appropriate control when the violation in responses to images was added as an extra critical factor to triggering a mismatch. Scrambled sequences inherently would result in much more response violations in the current design than the mismatch sequences, so, while these sequences still should not trigger a mnemonic mismatch, the response violations might have been just as efficient, if not more so, for triggering new encoding. Indeed, we did see a higher average novel recall rate for the scrambled condition than the mismatch, although this difference was not significant in our sample.

An additional factor that might have led to our lack of significant results was the increased demands resulting from the need to reply to the relationship judgments on each trial. To complete the task during the learning and mismatch phases of previous experiment participants needed to hold in mind a representation of the previously viewed image, evaluate the current image to that previously viewed item, and finally, keep a running tally of the number of times the relationship was met. The design of this experiment replaced the relatively easy task of keeping a single number in mind with the much more demanding need to respond before the next item. This increased task demand might have proven distracting to participants, and reduced the overall integrations of the items into a strong sequence memory. A final issue with the design of this experiment is that images, both those that made up the original sequence, and those that served as the novel items, were randomly assigned to sequences; whether

the novel image actually violated the predicted response was not controlled. Given these weaknesses, we cannot interpret these results much beyond acknowledging that any mismatch benefit is not a strong enough effect to overcome these issues.

Experiment 8: Set Novelty

As a control condition our scrambled sequence has many features to recommend it. By presenting items from a sequence out of order this condition matched the familiarity of the items presented with the critical novel items, without inducing the same mnemonic prediction expected from an intact sequence. However, one issue with this control condition is that while the presentation of a novel image does not represent a prediction violation, it does create a set mismatch. The novel item is unexpected in terms of the set of items that had been previously presented together. Furthermore, the presentation of items out of order is an order mismatch itself. This experiment attempts to control for these two factors in order to show the mismatch benefit more definitively.

Design

To eliminate set and order mismatches from the comparison condition, this version of the task design replaces the scrambled sequences condition with a shuffled condition. During the learning phase, sequences in the shuffled conditions no longer present fixed image- and order-sets, rather all images from the sequences binned into the shuffled condition are put into a pool, and for each block of the learning phase new sequences are created from this pool. Images are pulled from the pool of images to create unique sets of images, with no two images previously presented in the same sequence, for each learning block. Also, every item is seen in a different position within a sequence on each of its learning phase presentations. During the mismatch phase the

Experiment 8:

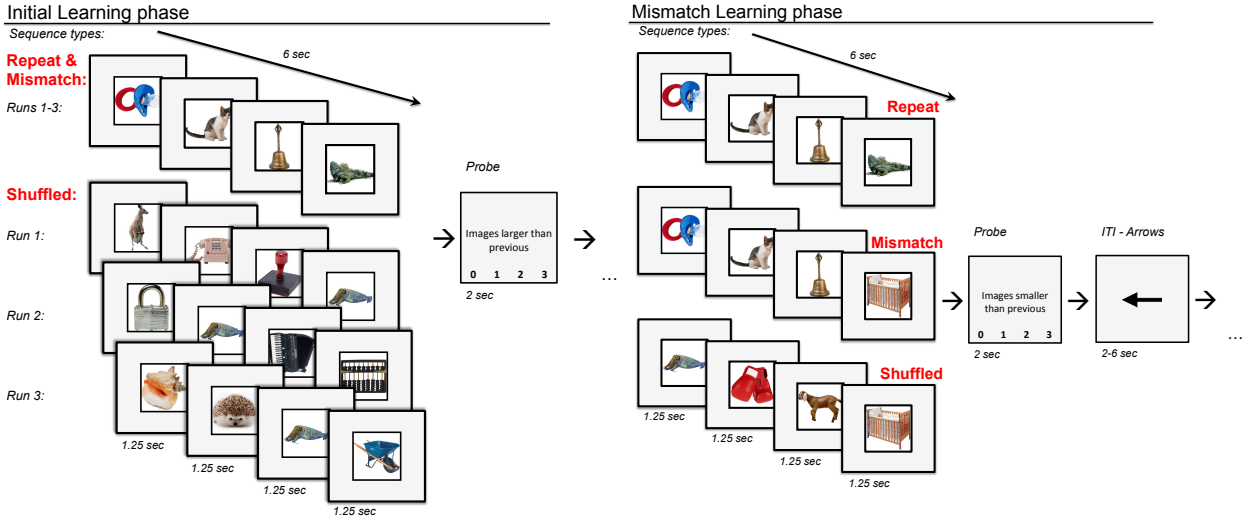


Figure 4.10: Experimental Design, part 3. Learning phases for experiment 8.

sequences in the shuffled condition again present a new unique set of images from the pool, but the last item in these sequences is a novel image. Figure 4.10 shows the design for this experiment.

Repeat and mismatch conditions still present the same sequences during the learning phase, and the mismatch condition still presents a new image in the final position during the mismatch phase, while the repeat sequences present no changes from the learning phase. This Experiment presents 95 sequences, 25 repeat, 35 mismatch, and 35 shuffled, and tests 285 individual items during the testing phase, a day after the learning and mismatch phase.

Participants

Twenty-one (15 female) participants completed the first version of the task ($mean_{age}$ = 20.77 years). Written consent was obtained from all participants in compliance with the UCLA Institutional Review Board, and all were compensated with course credits for their time. Three subjects preformed at or below chance on recall of all novel items, and were excluded. Data from three subjects was excluded due to

performances below chance on recall for all novel images. In total 18 (12 female) participants make up this dataset.

Results

Our results are contrary to our expectations that shuffling the images that make up the control sequences would result in less memory based expectations and its proposed benefit to encoding due to violations. We did not find any difference between the recall rates for novel images presented in mismatch sequences (*mean* = 42.90%, *SD* = 11.40%) compared to those in the shuffled condition (*mean* = 46.45%. *SD* = 9.57%; $t_{(17)}$ = 1.69, p = 0.11). Furthermore, when converted to novel costs (see Experiment 4), shuffled sequences (*mean* = -22.84%, *SD* = 14.58%) had a slight, though not significant, advantage over mismatches (*mean* = -28.78%, *SD* = 12.05%; $t_{(17)}$ = 1.966, p = 0.065). This advantage was not due to a significant difference in recall rates for non-novel images in these two conditions (mismatch: *mean* = 71.69%, *SD* = 11.70%; shuffled: *mean* = 69.29%, *SD* = 14.84%; $t_{(17)}$ = 1.24, p = 0.232). See Figure 4.11

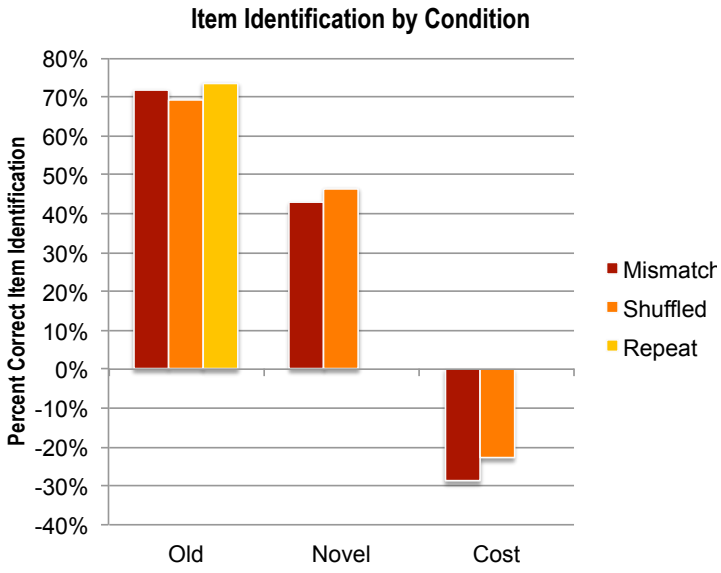


Figure 4.11: Results for Experiment 7. Novel items presented as the mismatch, shuffled condition as control. Both recall rates, and novel costs shown

Discussion

The shuffling of items to create a new set on each viewing, rather than serving as a better control by reducing the set violation of the scrambled condition, resulted in a slight reduction in novel costs for the items presented in this condition. A reduction in novel cost is an indication of stronger recall for novel images controlling for slight variations in overall recall by conditions. This benefit for the shuffled condition is unexpected based on the model of prediction violation as a trigger for new learning. This experiment highlights one of the critical difficulties with trying to isolate the mismatch benefit: while mnemonic mismatches might be a trigger for new learning, it is not the only such factor that can lead to improved encoding. These results may be an example of how task orientation can be a significant modulator of encoding. During both phases of this participants are required to make relationship judgments, the goal of which is to create a representation of the items as an ordered series. Set integration, therefor, is the ultimate goal of our task. The shuffled condition presented a completely novel set of images, and thus presented the most novelty in terms of the task goals. This task-oriented novelty, it would seem, created greater benefits for novel items than the mnemonic novelty resulting from the replacement of a single item of a sequence.

Meta-Analysis

Repeated Images Recall

In order to see the overall effect of repeated verse variable sequence presentations on item encoding we combined the results from Experiments 2-7. These experiments all presented sequences one or more times without changes during the learning phase. During the mismatch phase a portion were again presented without variation (repeats) while another portion presented some change to the sequence

(variants): either changes in sequence order, inclusion of novel images, or both. The memory for the items that made up the repeated sequences, when combining across all six experiments, was greater (*mean* = 69.83%, *SD* = 16.71%) than for the items that made up the variant sequences (*mean* = 68.14%, *SD* = 16.57%) ($t_{(120)}$=3.08, p = 0.003), Figure 4.12. These results support the conclusion of Experiment 3 that in our basic task design consistency, rather than variation, increased recall for items presented on multiple occasions.

Novel Images Recall

To examine the overall effect of otherwise consistent presentations of sequences on the recall for previously novel images we combined the results from Experiments 4-7. Each of these studies replaced one item of a sequence with a novel image during the mismatch phase. For sequences in the mismatch condition the rest of the original items were presented in the same order and position as during the learning phase, and in the scrambled condition these old items were presented out of order and position. When combining across all four experiments the novel items presented in the mismatch condition were recalled more (*mean* = 44.76%, *SD* = 13.25%), but not significantly so, than items in the scrambled condition (*mean* = 43.34%, *SD* = 14.05%) ($t_{(87)}$ = 1.098, p = 0.275).

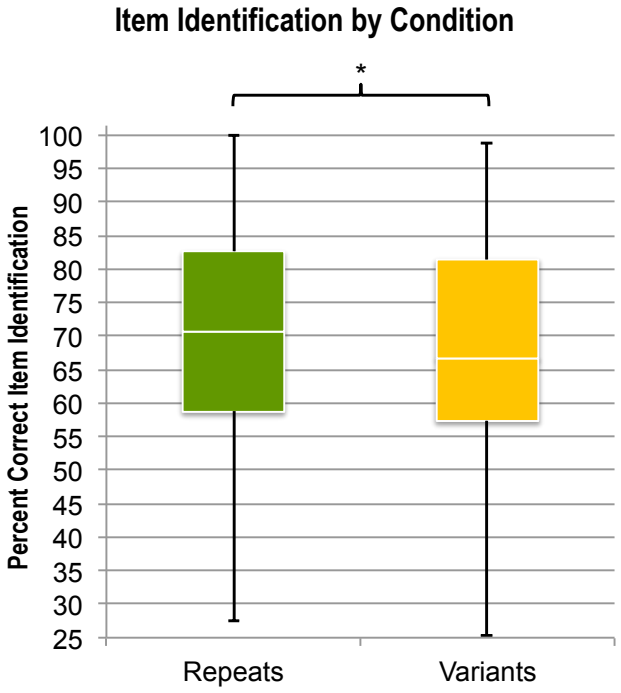


Figure 4.12: Results from Meta-analysis. All non-novel images in the repeat condition verse non-novels from all other conditions.

However, two of the experiments (6 and 7) may have unintentionally introduced a significant amount of overall novelty beyond the prediction violation we were interested in, see the discussion sections of each of these experiments for more details. When we exclude these two experiments, the difference between novel recall rates for mismatch (*mean* = 45.66%, *SD* = 13.54) and scrambled (*mean* = 42.93%, *SD* = 13.02%) conditions is still not significant, but we do find a mismatch benefit trend ($t_{(54)}$ = 1.69, p = 0.094). This trend provides a modest degree of support for the mismatch benefit seen in the final position of sequences in Experiment 4 and the similarly modest effect reported in Chen et al. (2015).

General Discussion

Here we conducted a number of experiments in order to evaluate a critical component of the model of mismatch detection systems within the hippocampus: the theory that this error signaling in the CA1 serves to trigger encoding of these unexpected materials (Lisman & Grace, 2005). Experiment 4 shows this hypothesized mismatch benefit for greater later recall of novel items that violated mnemonic predictions over similarly novel images whose presentation did not induce such a violation. Our results provide significant additional support to the conclusions drawn in Chen et al. (2015) regarding their own behavioral results, the only other extant study on this behavioral mismatch benefit. By replicating the mismatch benefit for subsequent recall in a four-alternative forced choice design, and creating the mismatch with a novel image we are able to conclude that the marked increase for recall of these images seen at the end of a sequence reflects much stronger encoding of these images as a direct result of the prediction violation during their only encounter prior to testing.

While from these behavioral results alone we cannot conclude that the recall benefit for mnemonic mismatches results from activity within the hippocampus, or any other brain region for that matter, when these results are evaluated in combination with previous functional imaging studies a larger story emerges. These studies find increased activity within the CA1 subfield of the hippocampus in response to mnemonic mismatches (Chen et al., 2015, 2011; Duncan et al., 2012; Kumaran & Maguire, 2006, 2007a, 2009), results consistent with a novelty detection role hypothesized based on the cytoarchitecture of this region (Amaral, 1993; Amaral & Witter, 1989; Kahn & Shohamy, 2013; Lisman & Grace, 2005). Such mnemonic novelty detection is thought to trigger a functional loop between the hippocampus and the ventral tegmental area (VTA), which serves to allow this detection signal to promote the encoding into long-term memory of behaviorally important information (Lisman & Grace, 2005). Given this model, we can infer that the presentation of novel images in place of anticipated items resulted in a novelty signal within the hippocampus, which, in turn, propagated through various nodes of the Hippocampal-VTA loop, ultimately triggering dopamine release in the hippocampus, boosting long term potentiation (LTP) processes (Shohamy & Adcock, 2010). The addition of our behavioral evidence indicating that mismatches boost encoding to the existing body of research evaluating this Hippocampal-VTA loop significantly strengthens our understanding of this model.

While the results of Experiment 4 align with the Hippocampal-VTA loop model, the results of the follow-up experiments failed to show anticipated improvements to the mismatch benefit in response to increased memory and prediction strengths, or the inclusion of task components. The results of Experiments 5-8, in fact, failed to replicate

the mismatch benefit in almost all cases, and the meta-analysis combining across all relevant studies showed no difference between the mismatch and scrambled conditions (although we did see a trend towards greater novel recall for the mismatch condition when we limited our meta-analysis to the studies that did not introduce extra novelty during the mismatch phase). As noted in the above discussions of each experiment, there were elements incorporated into each of our designs that, although intended to manipulate effects of interest, may have unintentionally interfered with our ability to observe the mismatch benefit. The fact that such small changes in experimental design can eliminate the effect would seem to indicate that it is not a particularly robust effect. The results of Experiment 3, which found greater subsequent memory for items presented in a consistent order in all viewings than those that swapped the order of the last two items, also indicated that any benefit triggered by mnemonic mismatches was not large enough to overcome the benefit resulting from consistency. The weak mismatch benefit, only a single percent increase in hit rate, reported by Chen et al. (2015) further emphasizes this point.

The susceptibility of the behavioral benefit to disruption or suppression by numerous other factors is, in many ways, not surprising. The downward arc of the Hippocampal-VTA loop is rather complex: the novelty signal generated in the CA_1 is propagated through the subiculum to accumbens neurons that are excited by numerous other brain regions (French & Totterdell, 2002), before continuing on to the ventral pallidum, and in turn the VTA. The VTA itself requires concurrent excitatory glutamatergic signals driven by limbic and prefrontal cortexes along with the disinhibition resulting from the novelty signal in order to induce a substantive release of

dopamine (Lodge & Grace, 2008), allowing this region to serve as a mediator of several systems (Oades & Halliday, 1987). The extensive number of inputs to this downward arc is thought to allow for significant modulation of novelty induced encoding by salience information and goal related motivations (Lisman & Grace, 2005).

Several studies, including the results of the study presented in Chapter III, have shown that the presentation of mnemonic mismatches in a task that requires explicit memory judgments causes the activity within the CA_1 subfield to increase in response to stimuli that match, rather than mismatch, previous expectations. The influence of task demands on the response within this critical node of the Hippocampal-VTA loop potentially reflects the end results of neuromodulatory inputs of the downward arc rather than the actual punitive mismatch response that would have triggered this loop initially (Lisman & Grace, 2005). That inputs from motivation and goal representing regions can potentially supersede the initial response of the CA_1 is testament to the influence of this information on this loop, and thus what is entered into long-term memory. Any experiment that hopes to demonstrate the mismatch benefit for subsequent recall needs to carefully control these motivational influences, and the failure to find effects in several of our experiments demonstrates the numerous difficulties in doing so.

The results of experiments 5 & 6 indicate that increasing the strength of the mismatch response does not necessarily lead to increased encoding of the novel material. Experiment 5, which employed a very similar design as Chen et al. (2015), which showed increased prediction violation signals within the hippocampus for stronger memories, failed to find a corresponding increase in behavioral benefits. While the novel items presented in our strong condition likely triggered a larger mismatch signal in the

CA_1, one or more of the salience or goal related motivations might have mediated the encoding influence of this violation signal. Experiment 6, likewise, did not find an increased benefit for items seen at the end of sequences, which would be expected to induce stronger predictions than earlier items. Although it is worth noting that the neural effect of this behavioral manipulation has not been evaluated, and may not, in fact, lead to greater hippocampal mismatch responses. Despite the lack of significant results from these two experiments, that the mismatch benefit was only found in items seen at the end of sequences in Experiment 4 might still indicate that increased prediction strength does lead to greater encoding. Future studies that are able to carefully control task motivations may yet uncover this strength benefit.

Given the multiple inputs to the Hippocampal-VTA loop from the PFC (French & Totterdell, 2002), and the disruptive effect of minor task changes had on our ability to detect the mismatch benefit, there is a large scope for future studies into task-focused effects on the mismatch benefit. Our Experiment 7 attempted to boost the mismatch benefit by incorporating task violation elements, but we did not adequately account for how these task demands would affect our control condition, so it is difficult to draw many conclusions from this experiment, but the results of Experiment 8 would indicate an important role of novelty relative to task orientations. In every experiment conducted here our task during the learning phase asked relationship judgments in order to integrate the items into an ordered set representation; the goal of this task, during both the learning and mismatch phases of Experiment 8, was set integration. When framing our task in this light it is clear that our shuffled condition would have presented the highest degree of task-oriented novelty; each shuffled sequence presented a wholly

novel set of images requiring integration. The inclusion of novel images in mismatch sequences also presents a degree of set novelty, but to a much smaller extent. The findings of Experiment 8 would indicate that this goal-oriented novel set integration might results in just as much benefit to encoding as prediction violations.

The other major result of this series of experiments is that recall for images seen multiple times prior to testing was greatest when these items had been presented consistently through out all previous presentations. A meta-analysis of our seven experiments that tested participants recall for the items within sequences finds conclusively that when some degree of variation in the sequence, whether resulting from the inclusion of a novel image, or the rearranging of items order, occurred during the testing phase the non-novel items that made up that sequence were recalled at lower rates than for items in sequences with no such variations. While this result was unexpected based on the hypotheses generated from the mismatch detection model, they are not wholly unprecedented in behavioral examinations of how multiple presentations of materials can improve memory. One consistent finding in this field is that spaced repetition of materials can lead to marked increases for subsequent recall over blocked repetition (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). However, when the to-be-learned material varies between the well-spaced presentations later recall can suffer (Appleton-Knapp, Bjork, & Wickens, 2005). The variation between learning events, combined with the prolonged time between presentations may reduce the retrieval of references to the earlier presentations, the proposed mechanism behind the benefit of spaced learning (Storm, Bjork, & Storm, 2010). Furthermore, greater similarity between the disperse neural response pattern as measured with functional

MRI has been shown to predict greater subsequent recall (Xue et al., 2010). In our tasks, which presented sequences spaced between several learning blocks and phases, the variation in mismatch and scrambled sequences seem to have interfered with the recall of previous presentations of these sequences, and thus reduced their overall encoding.

Conclusion

Here we show behavioral evidence for the proposed theory that detection of mnemonic violations within the hippocampus (Chen et al., 2015, 2011; Duncan et al., 2012; Kumaran & Maguire, 2006, 2007a, 2009) acts as a trigger, inducing increased encoding of these memory-based mismatches (Lisman & Grace, 2005). We show increased subsequent recall for novel items that replaced expected items at the end of well-learned sequences, likely resulting from increased LTP promoted by the release of dopamine (Shohamy & Adcock, 2010), the end result of the violation elicited Hippocampal-VTA loop (Lisman & Grace, 2005). Follow-up experiments identified some of the difficulties inherent to behavioral isolations of specific effects resulting from elements of this complex hippocampal-VTA network. This study lays a strong foundation for future projects that seek to further examine this mismatch induced recall benefit.

CHAPTER 5

General Discussion

This dissertation has presented three research projects examining how hippocampal functions, particularly those assigned to specific subfields, are modulated by task and context demands. Each of these projects resulted in several interesting findings, and provides further support for and insight into the theories describing mechanisms underlying memory in humans. In the study presented in Chapter 2, we were able to demonstrate dissociable response profiles within hippocampal subfields as a function of probe novelty and associative context. The results of this project replicate previous work showing a dissociation between discrimination processes in the CA_3DG and generalization in the CA1 (Bakker, Kirwan, Miller, & Stark, 2008; Lacy, Yassa, Stark, Muftuler, & Stark, 2011). The computational difference between these two subfields has previously only been seen in implicit learning paradigms, evident through repetition suppression effects, whereas our results demonstrate this dissociation in the context of match enhancement effects. The persistence of this dissociation in an explicit recall task has several important implications regarding the attribution of differential lure sensitivity of the CA_3DG and CA_1 subfields to pattern separation and pattern completion mechanisms. Pattern separation is an encoding processing (Hunsaker & Kesner, 2013), and so the discrimination seen in the CA_3DG during our explicit recall task is not likely a reflection of pattern separation.

The results presented in Chapter 2 also replicated the increased activity within the CA_1 to matching conditions when subjects were explicitly tasked with evaluating (Dudukovic et al., 2011; Duncan et al., 2009, 2012; Hannula & Ranganath, 2008),

although this match enhancement was seen in the CA_3DG, unlike previous studies. While both CA_3DG and CA_1 responded more to correctly paired matches than novels, their response to previously seen, but mispaired images differed. CA_3DG discriminated old visual stimuli from new stimuli, irrespective of the task-relevance context, while CA_1 responded to incorrectly paired repeats as it would to novel. This dissociation indicate that the CA_1 is the subfield particularly sensitive to motivational and goal information.

Further evidence for the CA_1 playing an important role in the integration of top-down, goal-directed information with memory functions was seen in the study presented in Chapter 3, where only the CA_1 showed greater functional coupling with the VTA for high value cues than low value ones. This increased coupling likely reflects the neuromodulatory inputs from the downward arc of the hippocampal-VTA loop (Lisman & Grace, 2005). The CA_3DG did not display coupling with reward regions, rather different reward values modulated encoding processes in this region. Our findings indicate that items assigned a low reward value are putatively processed within the medial temporal lobe, particular the right CA_3DG, in a less consistent manner than higher value items at encoding. Within the CA_3DG, when compared to the baseline similarity between the representation of target faces at encoding and novel faces at retrieval the similarity between target faces assigned low values at encoding, and the same face presented again during retrieval was reduced. This decrease in consistency likely reflects reduced pattern separation processes in the CA_3DG.

Previous studies had suggested that reward values are represented within the hippocampus (Wolosin et al., 2013), however, our results indicate otherwise. Only in the fusiform cortex showed increased within values representational similarity at

encoding for items assigned high value then for those assigned low value. No other MTL cortex or hippocampal subfields show a difference in representational similarity as a function of reward, indicating that these regions do not represent reward. Furthermore, no MTL regions, including the fusiform, showed differences in similarities due to reward states at retrieval, connoting a lack of reinstatement of reward value. While disputing the findings of Wolosin et al. (2013), these results align with other studies proposing that the processing of information within the hippocampus is domain agnostic (Azab et al., 2014; Huffman & Stark, 2014). The results of this project demonstrate how the assignment of a reward value to an item can modulate hippocampal processing.

Finally, Chapter 4 attempted to examine the behavioral results of the mismatch response observed in the CA_1. The model of mismatch detection systems within the hippocampus posits that the punitive error signaling in the CA_1 serves to trigger encoding of the unexpected material (Kumaran & Maguire, 2009; Lisman & Grace, 2005). We conducted a number of experiments in order to detect this theorized recall benefit for mismatching stimuli. In our experiment 4 we demonstrated increased recall rates on a four-alternative forced choice recognition task for novel items that violated mnemonic predictions over similarly novel images whose presentation did not induce such a violation. Our findings aligns with those of Chen et al. (2015), and our design, which created the mismatch with a novel image encountered only once prior to testing, allows us to conclude that the marked increase for recall of these images seen at the end of a sequence reflects stronger encoding of these images as a direct result of the prediction violation. When evaluated in combination with previous functional imaging

studies our behavioral results support Hippocampal-VTA loop model (Lisman & Grace, 2005). In this model, the novelty detection in the CA1 shown in these studies (Chen et al., 2015, 2011; Duncan et al., 2012) triggers new learning via dopamine release by the VTA. Our results likely reflect this boosted encoding, and the addition of our behavioral evidence to the existing body of research evaluating this Hippocampal-VTA loop significantly strengthens our understanding of this model.

The follow-up experiments to Experiment 4 lay groundwork for examining how other factors, such as encoding and prediction strengths might influence this behavioral mismatch benefit. The failure of these studies to replicate the mismatch benefit highlights the susceptibility of the effect to disruption or suppression by numerous other factors. This susceptibility and the modest size of the benefit seen in Chen et al. (2015), imply that this effect is not robust. The Hippocampal-VTA loop thought to drive this effect includes multiple inputs from several regions, allowing for the modulation of encoding by task and motivational information (Lisman & Grace, 2005). Our experiments demonstrate how slight manipulations in task demands can have large effects on the behavioral outputs of this loop. Future studies on this behavioral mismatch benefit will need to carefully consider and control for these susceptibilities. Another result seen in the experiments presented in Chapter 4 is that recall for images viewed multiple times prior to testing was greatest when these items had been presented consistently throughout all previous presentations. Although unexpected in terms of the mismatch detection model, this results reconciles with findings that variations in well-spaced presentations can reduce encoding (Appleton-Knapp et al.,

2005), and more consistent mental representation of materials predicts greater subsequent recall (Xue et al., 2010).

In conclusion, the work presented in this dissertation adds to a growing body of literature describing diverging computational processes within hippocampal subfields. Furthermore, we have shown that task demands and reward values differentially modulate these regions. Finally, we have demonstrated how mnemonic content can trigger a shift to encoding when present conditions deviate from the expected. Future research expanding on these results will provide valuable insights to questions of how hippocampal functions allow for flexible goal-modulated encoding and retrieval processes.

References

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. E. (2006). Reward-Motivated Learning: Mesolimbic Activation Precedes Memory Formation. *Neuron*, *50*(3), 507–517. http://doi.org/10.1016/j.neuron.2006.03.036

Allen, T. A., & Fortin, N. J. (2013). The evolution of episodic memory. *Proceedings of the National Academy of Sciences*, *110*(Supplement_2), 10379–10386. http://doi.org/10.1073/pnas.1301199110

Allen, T. A., Morris, A. M., Stark, S. M., Fortin, N. J., & Stark, C. E. L. (2015). Memory for sequences of events impaired in typical aging. *Learning & Memory*, *22*(3), 138–148. http://doi.org/10.1101/lm.036301.114

Amaral, D. G. (1993). Emerging principles of intrinsic hippocampal organization. *Current Opinion in Neurobiology*, *3*(2), 225–229.

Amaral, D. G., & Witter, M. P. (1989). The three-dimensional organization of the hippocampal formation: A review of anatomical data. *Neuroscience*, *31*(3), 571–591. http://doi.org/10.1016/0306-4522(89)90424-7

Andrews-Hanna, J. R. (2012). The Brain's Default Network and Its Adaptive Role in Internal Mentation. *The Neuroscientist*, *18*(3), 251–270. http://doi.org/10.1177/1073858411403316

Appleton-Knapp, S. L., Bjork, R. A., & Wickens, T. D. (2005). Examining the Spacing Effect in Advertising: Encoding Variability, Retrieval Processes, and Their Interaction. *Journal of Consumer Research*, *32*(2), 266–276. http://doi.org/10.1086/432236

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, *54*(3), 2033–2044. http://doi.org/10.1016/j.neuroimage.2010.09.025

Avants, B. B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J., & Gee, J. C. (2010). The optimal template effect in hippocampus studies of diseased populations. *NeuroImage*, *49*(3), 2457–2466. http://doi.org/10.1016/j.neuroimage.2009.09.062

Azab, M., Stark, S. M., & Stark, C. E. L. (2014). Contributions of human hippocampal subfields to spatial and temporal pattern separation: Domain Agnostic Role for the Dentate Gyrus. *Hippocampus*, *24*(3), 293–302. http://doi.org/10.1002/hipo.22223

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. L. (2008). Pattern Separation in the Human Hippocampal CA3 and Dentate Gyrus. *Science*, *319*(5870), 1640–1642. http://doi.org/10.1126/science.1152882

Bangasser, D. A. (2006). Trace Conditioning and the Hippocampus: The Importance of Contiguity. *Journal of Neuroscience*, *26*(34), 8702–8706. http://doi.org/10.1523/JNEUROSCI.1742-06.2006

Benjamin, A. S. (2007). Memory is More than just Remembering: Strategic Control of Encoding, Accessing Memory, and Making Decisions. In A. S. B. and B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 48, pp. 175–223). Academic Press. Retrieved from http://www.sciencedirect.com/science/article/pii/S0079742107480057

Bonnici, H. M., Chadwick, M. J., Lutti, A., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2012). Detecting Representations of Recent and Remote Autobiographical Memories in vmPFC and Hippocampus. *Journal of Neuroscience*, *32*(47), 16982–16991. http://doi.org/10.1523/JNEUROSCI.2475-12.2012

Buckner, R. L. (2010). The role of the hippocampus in prediction and imagination. *Annual Review of Psychology*, *61*, 27–48, C1–8. http://doi.org/10.1146/annurev.psych.60.110707.163508

Carr, V. A., Rissman, J., & Wagner, A. D. (2010). Imaging the Human Medial Temporal Lobe with High-Resolution fMRI. *Neuron*, *65*(3), 298–308. http://doi.org/10.1016/j.neuron.2009.12.022

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354–380. http://doi.org/10.1037/0033-2909.132.3.354

Chadwick, M. J., Bonnici, H. M., & Maguire, E. A. (2012). Decoding information in the human hippocampus: A user's guide. *Neuropsychologia*, *50*(13), 3107–3121. http://doi.org/10.1016/j.neuropsychologia.2012.07.007

Chadwick, M. J., Hassabis, D., & Maguire, E. A. (2011). Decoding overlapping memories in the medial temporal lobes using high-resolution fMRI. *Learning & Memory*, *18*(12), 742–746. http://doi.org/10.1101/lm.023671.111

Chadwick, M. J., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2010). Decoding Individual Episodic Memory Traces in the Human Hippocampus. *Current Biology*, *20*(6), 544–547. http://doi.org/10.1016/j.cub.2010.01.053

Chen, J., Cook, P. A., & Wagner, A. D. (2015). Prediction strength modulates responses in human area CA1 to sequence violations. *Journal of Neurophysiology*, *114*(2), 1227–1238. http://doi.org/10.1152/jn.00149.2015

Chen, J., Dastjerdi, M., Foster, B. L., LaRocque, K. F., Rauschecker, A. M., Parvizi, J., & Wagner, A. D. (2013). Human hippocampal increases in low-frequency power during associative prediction violations. *Neuropsychologia*, *51*(12), 2344–2351. http://doi.org/10.1016/j.neuropsychologia.2013.03.019

Chen, J., Olsen, R. K., Preston, A. R., Glover, G. H., & Wagner, A. D. (2011). Associative retrieval processes in the human medial temporal lobe: hippocampal retrieval success and CA1 mismatch detection. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *18*(8), 523–528. http://doi.org/10.1101/lm.2135211

Cohen, M. S., Rissman, J., Suthana, N. A., Castel, A. D., & Knowlton, B. J. (2014). Value-based modulation of memory encoding involves strategic engagement of fronto-temporal semantic processing regions. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 578–592. http://doi.org/10.3758/s13415-014-0275-x

Deuker, L., Doeller, C. F., Fell, J., & Axmacher, N. (2014). Human neuroimaging studies on the hippocampal CA3 region - integrating evidence for pattern separation and completion. *Frontiers in Cellular Neuroscience*, *8*. http://doi.org/10.3389/fncel.2014.00064

Dudukovic, N. M., Preston, A. R., Archie, J. J., Glover, G. H., & Wagner, A. D. (2011). High-resolution fMRI reveals match enhancement and attentional modulation in the human medial temporal lobe. *Journal of Cognitive Neuroscience*, *23*(3), 670–682.

Duncan, K., Curtis, C., & Davachi, L. (2009). Distinct Memory Signatures in the Hippocampus: Intentional States Distinguish Match and Mismatch Enhancement Signals. *Journal of Neuroscience*, *29*(1), 131–139. http://doi.org/10.1523/JNEUROSCI.2998-08.2009

Duncan, K., Ketz, N., Inati, S. J., & Davachi, L. (2012). Evidence for area CA1 as a match/mismatch detector: A high-resolution fMRI study of the human hippocampus. *Hippocampus*, *22*(3), 389–398. http://doi.org/10.1002/hipo.20933

Duvernoy, H. M., Cattin, F., & Risold, P.-Y. (2013). *The Human Hippocampus*. Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://link.springer.com/10.1007/978-3-642-33603-4

Eichenbaum, H. (2000). Hippocampus: Mapping or memory? *Current Biology*, *10*(21), R785–R787. http://doi.org/10.1016/S0960-9822(00)00763-6

Eichenbaum, H. (2013). Memory on time. *Trends in Cognitive Sciences*, *17*(2), 81–88. http://doi.org/10.1016/j.tics.2012.12.007

Ezzyat, Y., & Davachi, L. (2011). What Constitutes an Episode in Episodic Memory? *Psychological Science*, *22*(2), 243–252. http://doi.org/10.1177/0956797610393742

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47.

Frankó, E., Insausti, A. M., Artacho-Pérula, E., Insausti, R., & Chavoix, C. (2014). Identification of the human medial temporal lobe regions on magnetic resonance images: Human Medial Temporal Lobe Landmarks. *Human Brain Mapping*, *35*(1), 248–256. http://doi.org/10.1002/hbm.22170

French, S. J., & Totterdell, S. (2002). Hippocampal and prefrontal cortical inputs monosynaptically converge with individual projection neurons of the nucleus accumbens. *The Journal of Comparative Neurology*, *446*(2), 151–165. http://doi.org/10.1002/cne.10191

Gasbarri, A., Packard, M. G., Campana, E., & Pacitti, C. (1994). Anterograde and retrograde tracing of projections from the ventral tegmental area to the hippocampal formation in the rat. *Brain Research Bulletin*, *33*(4), 445–452. http://doi.org/10.1016/0361-9230(94)90288-7

Hannula, D. E., & Ranganath, C. (2008). Medial Temporal Lobe Activity Predicts Successful Relational Memory Binding. *Journal of Neuroscience*, *28*(1), 116–124. http://doi.org/10.1523/JNEUROSCI.3086-07.2008

Hassabis, D., Chu, C., Rees, G., Weiskopf, N., Molyneux, P. D., & Maguire, E. A. (2009). Decoding Neuronal Ensembles in the Human Hippocampus. *Current Biology*, *19*(7), 546–554. http://doi.org/10.1016/j.cub.2009.02.033

Hassabis, D., Kumaran, D., Vann, S. D., & Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, *104*(5), 1726–1731. http://doi.org/10.1073/pnas.0610561104

Huffman, D. J., & Stark, C. E. L. (2014). Multivariate pattern analysis of the human medial temporal lobe revealed representationally categorical cortex and representationally agnostic hippocampus. *Hippocampus*, *24*(11), 1394–1403. http://doi.org/10.1002/hipo.22321

Hunsaker, M. R., & Kesner, R. P. (2013). The operation of pattern separation and pattern completion processes associated with different attributes or domains of memory. *Neuroscience & Biobehavioral Reviews*, *37*(1), 36–58. http://doi.org/10.1016/j.neubiorev.2012.09.014

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review*, *2*(1), 105–112.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. http://doi.org/10.1006/nimg.2002.1132

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. http://doi.org/10.1016/j.neuroimage.2011.09.015

Kahn, I., Andrews-Hanna, J. R., Vincent, J. L., Snyder, A. Z., & Buckner, R. L. (2008). Distinct Cortical Anatomy Linked to Subregions of the Medial Temporal Lobe Revealed by Intrinsic Functional Connectivity. *Journal of Neurophysiology*, *100*(1), 129–139. http://doi.org/10.1152/jn.00077.2008

Kahn, I., & Shohamy, D. (2013). Intrinsic connectivity between the hippocampus, nucleus accumbens, and ventral tegmental area in humans. *Hippocampus*, *23*(3), 187–192. http://doi.org/10.1002/hipo.22077

Kesner, R. P. (2013a). An analysis of the dentate gyrus function. *Behavioural Brain Research*, *254*, 1–7. http://doi.org/10.1016/j.bbr.2013.01.012

Kesner, R. P. (2013b). A process analysis of the CA3 subregion of the hippocampus. *Frontiers in Cellular Neuroscience*, *7*. http://doi.org/10.3389/fncel.2013.00078

Kesner, R. P. (2013c). Role of the hippocampus in mediating interference as measured by pattern separation processes. *Behavioural Processes*, *93*, 148–154. http://doi.org/10.1016/j.beproc.2012.09.018

Kirwan, C. B., Jones, C. K., Miller, M. I., & Stark, C. E. L. (2007). High-resolution fMRI investigation of the medial temporal lobe. *Human Brain Mapping*, *28*(10), 959–966. http://doi.org/10.1002/hbm.20331

Kirwan, C. B., & Stark, C. E. L. (2007). Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learning & Memory*, *14*(9), 625–633. http://doi.org/10.1101/lm.663507

Kishiyama, M. M., Yonelinas, A. P., & Lazzara, M. M. (2004). The von Restorff effect in amnesia: the contribution of the hippocampal system to novelty-related memory enhancements. *Journal of Cognitive Neuroscience*, *16*(1), 15–23. http://doi.org/10.1162/089892904322755511

Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*. http://doi.org/10.3389/neuro.06.004.2008

Kuhl, B. A., Rissman, J., & Wagner, A. D. (2012). Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia*, *50*(4), 458–469. http://doi.org/10.1016/j.neuropsychologia.2011.09.002

Kumaran, D., & Maguire, E. A. (2006). An unexpected sequence of events: mismatch detection in the human hippocampus. *PLoS Biol*, *4*(12), e424.

Kumaran, D., & Maguire, E. A. (2007a). Match Mismatch Processes Underlie Human Hippocampal Responses to Associative Novelty. *Journal of Neuroscience*, *27*(32), 8517–8524. http://doi.org/10.1523/JNEUROSCI.1677-07.2007

Kumaran, D., & Maguire, E. A. (2007b). Which computational mechanisms operate in the hippocampus during novelty detection? *Hippocampus*, *17*(9), 735–748. http://doi.org/10.1002/hipo.20326

Kumaran, D., & Maguire, E. A. (2009). Novelty signals: a window into hippocampal information processing. *Trends in Cognitive Sciences*, *13*(2), 47–54. http://doi.org/10.1016/j.tics.2008.11.004

Lacy, J. W., & Stark, C. E. L. (2012). Intrinsic functional connectivity of the human medial temporal lobe suggests a distinction between adjacent MTL cortices and hippocampus. *Hippocampus*, *22*(12), 2290–2302. http://doi.org/10.1002/hipo.22047

Lacy, J. W., Yassa, M. A., Stark, S. M., Muftuler, L. T., & Stark, C. E. L. (2011). Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learning & Memory*, *18*(1), 15–18. http://doi.org/10.1101/lm.1971111

Lambrey, S., Amorim, M.-A., Samson, S., Noulhiane, M., Hasboun, D., Dupont, S., … Berthoz, A. (2008). Distinct visual perspective-taking strategies involve the left and right medial temporal lobe structures differently. *Brain*, *131*(2), 523–534. http://doi.org/10.1093/brain/awm317

Larkman, D. J., & Nunes, R. G. (2007). Parallel magnetic resonance imaging. *Physics in Medicine and Biology*, *52*(7), R15–55. http://doi.org/10.1088/0031-9155/52/7/R01

LaRocque, K. F., Smith, M. E., Carr, V. A., Witthoft, N., Grill-Spector, K., & Wagner, A. D. (2013). Global Similarity and Pattern Separation in the Human Medial Temporal Lobe Predict Subsequent Memory. *Journal of Neuroscience*, *33*(13), 5466–5474. http://doi.org/10.1523/JNEUROSCI.4293-12.2013

Leutgeb, J. K., Leutgeb, S., Treves, A., Meyer, R., Barnes, C. A., McNaughton, B. L., … Moser, E. I. (2005). Progressive Transformation of Hippocampal Neuronal Representations in “Morphed” Environments. *Neuron*, *48*(2), 345–358. http://doi.org/10.1016/j.neuron.2005.09.007

Leutgeb, S., & Leutgeb, J. K. (2007). Pattern separation, pattern completion, and new neuronal codes within a continuous CA3 map. *Learning & Memory*, *14*(11), 745–757. http://doi.org/10.1101/lm.703907

Liang, J. C., & Preston, A. R. (2015). Medial Temporal Lobe Subregional Function in Human Episodic Memory. In D. R. Addis, M. Barense, & A. Duarte (Eds.), *The Wiley Handbook on the Cognitive Neuroscience of Memory* (pp. 108–130). John Wiley & Sons, Ltd. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/9781118332634.ch6/summary

Libby, L. A., Ekstrom, A. D., Ragland, J. D., & Ranganath, C. (2012). Differential Connectivity of Perirhinal and Parahippocampal Cortices within Human Hippocampal Subregions Revealed by High-Resolution Functional Imaging. *Journal of Neuroscience*, *32*(19), 6550–6560. http://doi.org/10.1523/JNEUROSCI.3711-11.2012

Lisman, J. E., & Grace, A. A. (2005). The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. *Neuron*, *46*(5), 703–713. http://doi.org/10.1016/j.neuron.2005.05.002

Lodge, D. J., & Grace, A. A. (2008). Hippocampal dysfunction and disruption of dopamine system regulation in an animal model of schizophrenia. *Neurotoxicity Research*, *14*(2-3), 97–104. http://doi.org/10.1007/BF03033801

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, *71*(4), 737–749. http://doi.org/10.1016/j.neuron.2011.07.012

Milian, M., Zeltner, L., Klamer, S., Klose, U., Rona, S., & Erb, M. (2013). BOLD Signal in memory paradigms in hippocampal region depends on echo time. *Journal of Magnetic Resonance Imaging*, *37*(5), 1064–1071. http://doi.org/10.1002/jmri.23888

Moser, E. I., Kropff, E., & Moser, M.-B. (2008). Place Cells, Grid Cells, and the Brain's Spatial Representation System. *Annual Review of Neuroscience*, *31*(1), 69–89. http://doi.org/10.1146/annurev.neuro.31.061307.090723

Mullally, S. L., & Maguire, E. A. (2013). Memory, Imagination, and Predicting the Future A Common Brain Mechanism? *The Neuroscientist*, 1073858413495091.

Mumford, J. A. (2007). A guide to calculating percent change with featquery. Retrieved from http://mumford.bol.ucla.edu/perchange_guide.pdf

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification

analyses. *NeuroImage*, *59*(3), 2636–2643. http://doi.org/10.1016/j.neuroimage.2011.08.076

Murty, V. P., & Adcock, R. A. (2014). Enriched Encoding: Reward Motivation Organizes Cortical Networks for Hippocampal Detection of Unexpected Events. *Cerebral Cortex*, *24*(8), 2160–2168. http://doi.org/10.1093/cercor/bht063

Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., & Sederberg, P. B. (2015). Human hippocampus represents space and time during retrieval of real-world memories. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(35), 11078–11083. http://doi.org/10.1073/pnas.1507104112

Oades, R. D., & Halliday, G. M. (1987). Ventral tegmental (A10) system: neurobiology. 1. Anatomy and connectivity. *Brain Research Reviews*, *12*(2), 117–165. http://doi.org/10.1016/0165-0173(87)90011-7

O'Reilly, J. X., Woolrich, M. W., Behrens, T. E. J., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, *7*(5), 604–609. http://doi.org/10.1093/scan/nss055

Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*, *2*(1), 67–70. http://doi.org/10.1093/scan/nsm006

Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science (New York, N.Y.)*, *310*(5756), 1963–1966. http://doi.org/10.1126/science.1117645

Poppenk, J., Evensmoen, H. R., Moscovitch, M., & Nadel, L. (2013). Long-axis specialization of the human hippocampus. *Trends in Cognitive Sciences*, *17*(5), 230–240. http://doi.org/10.1016/j.tics.2013.03.005

Poppenk, J., & Moscovitch, M. (2011). A Hippocampal Marker of Recollection Memory Ability among Healthy Young Adults: Contributions of Posterior and Anterior Segments. *Neuron*, *72*(6), 931–937. http://doi.org/10.1016/j.neuron.2011.10.014

Preston, A. R., & Eichenbaum, H. (2013). Interplay of Hippocampus and Prefrontal Cortex in Memory. *Current Biology*, *23*(17), R764–R773. http://doi.org/10.1016/j.cub.2013.05.041

Ranganath, C., Yonelinas, A. P., Cohen, M. X., Dy, C. J., Tom, S. M., & D'Esposito, M. (2004). Dissociable correlates of recollection and familiarity within the medial temporal lobes. *Neuropsychologia*, *42*(1), 2–13.

Reas, E. T., & Brewer, J. B. (2013). Effortful retrieval reduces hippocampal activity and impairs incidental encoding. *Hippocampus*, *23*(5), 367–379. http://doi.org/10.1002/hipo.22096

Rissman, J., Gazzaley, A., & D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, *23*(2), 752–763. http://doi.org/10.1016/j.neuroimage.2004.06.035

Rissman, J., & Wagner, A. D. (2012). Distributed Representations in Memory: Insights from Functional Brain Imaging. *Annual Review of Psychology*, *63*(1), 101–128. http://doi.org/10.1146/annurev-psych-120710-100344

Ritchey, M., Wing, E. A., LaBar, K. S., & Cabeza, R. (2013). Neural Similarity Between Encoding and Retrieval is Related to Memory Via Hippocampal Interactions. *Cerebral Cortex*, *23*(12), 2818–2828. http://doi.org/10.1093/cercor/bhs258

Rotello, C. M., & Heit, E. (2000). Associative recognition: A case of recall-to-reject processing. *Memory & Cognition*, *28*(6), 907–922.

Samson, Y., Wu, J. J., Friedman, A. H., & Davis, J. N. (1990). Catecholaminergic innervation of the hippocampus in the cynomolgus monkey. *The Journal of Comparative Neurology*, *298*(2), 250–263. http://doi.org/10.1002/cne.902980209

Schomaker, J., Berendse, H. W., Foncke, E. M. J., van der Werf, Y. D., van den Heuvel, O. A., Theeuwes, J., & Meeter, M. (2014). Novelty processing and memory formation in Parkinson's disease. *Neuropsychologia*, *62*, 124–136. http://doi.org/10.1016/j.neuropsychologia.2014.07.016

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11–21.

Segal, S. K., Stark, S. M., Kattan, D., Stark, C. E., & Yassa, M. A. (2012). Norepinephrine-mediated emotional arousal facilitates subsequent pattern separation. *Neurobiology of Learning and Memory*, *97*(4), 465–469. http://doi.org/10.1016/j.nlm.2012.03.010

Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*, *14*(10), 464–472. http://doi.org/10.1016/j.tics.2010.08.002

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, *4*(8), 637–648. http://doi.org/10.1038/nrn1178

Spiers, H. J., & Maguire, E. A. (2007). The neuroscience of remote spatial memory: A tale of two cities. *Neuroscience*, *149*(1), 7–27. http://doi.org/10.1016/j.neuroscience.2007.06.056

Squire, L. R. (2009a). Memory and Brain Systems: 1969-2009. *Journal of Neuroscience*, *29*(41), 12711–12716. http://doi.org/10.1523/JNEUROSCI.3575-09.2009

Squire, L. R. (2009b). The Legacy of Patient H.M. for Neuroscience. *Neuron*, *61*(1), 6–9. http://doi.org/10.1016/j.neuron.2008.12.023

Staresina, B. P., & Davachi, L. (2009). Mind the Gap: Binding Experiences across Space and Time in the Human Hippocampus. *Neuron*, *63*(2), 267–276. http://doi.org/10.1016/j.neuron.2009.06.024

Stark, C. E. L., & Squire, L. R. (2000). Functional magnetic resonance imaging (fMRI) activity in the hippocampal region during recognition memory. *The Journal of Neuroscience*, *20*(20), 7776–7781.

Stark, C. E. L., & Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences*, *98*(22), 12760–12766. http://doi.org/10.1073/pnas.221462998

St. Jacques, P., Rubin, D. C., LaBar, K. S., & Cabeza, R. (2008). The short and long of it: Neural correlates of temporal-order memory for autobiographical events. *Journal of Cognitive Neuroscience*, *20*(7), 1327–1341. http://doi.org/10.1162/jocn.2008.20091

Storm, B. C., Bjork, R. A., & Storm, J. C. (2010). Optimizing retrieval as a learning event: When and why expanding retrieval practice enhances long-term retention. *Memory & Cognition*, *38*(2), 244–253. http://doi.org/10.3758/MC.38.2.244

Tsukiura, T., Sekiguchi, A., Yomogida, Y., Nakagawa, S., Shigemune, Y., Kambara, T., … Kawashima, R. (2010). Effects of Aging on Hippocampal and Anterior Temporal Activations during Successful Retrieval of Memory for Face–Name Associations. *Journal of Cognitive Neuroscience*, *23*(1), 200–213. http://doi.org/10.1162/jocn.2010.21476

Tubridy, S., & Davachi, L. (2011). Medial Temporal Lobe Contributions to Episodic Sequence Encoding. *Cerebral Cortex*, *21*(2), 272–280. http://doi.org/10.1093/cercor/bhq092

van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211–219. http://doi.org/10.1016/j.tins.2012.02.001

Vazdarjanova, A., & Guzowski, J. F. (2004). Differences in Hippocampal Neuronal Population Responses to Modifications of an Environmental Context: Evidence for Distinct, Yet Complementary, Functions of CA3 and CA1 Ensembles. *The Journal of Neuroscience*, *24*(29), 6489–6496. http://doi.org/10.1523/JNEUROSCI.0350-04.2004

Vinogradova, O. S. (2001). Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus*, *11*(5), 578–598.

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychologische Forschung*, *18*(1), 299–342. http://doi.org/10.1007/BF02409636

Wing, E. A., Ritchey, M., & Cabeza, R. (2015). Reinstatement of Individual Past Events Revealed by the Similarity of Distributed Activation Patterns during Encoding and Retrieval. *Journal of Cognitive Neuroscience*, *27*(4), 679–691. http://doi.org/10.1162/jocn_a_00740

Winterburn, J. L., Pruessner, J. C., Chavez, S., Schira, M. M., Lobaugh, N. J., Voineskos, A. N., & Chakravarty, M. M. (2013). A novel in vivo atlas of human hippocampal subfields using high-resolution 3T magnetic resonance imaging. *NeuroImage*, *74*, 254–265. http://doi.org/10.1016/j.neuroimage.2013.02.003

Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2012). Reward Modulation of Hippocampal Subfield Activation during Successful Associative Encoding and Retrieval. *Journal of Cognitive Neuroscience*, *24*(7), 1532–1547. http://doi.org/10.1162/jocn_a_00237

Wolosin, S. M., Zeithamova, D., & Preston, A. R. (2013). Distributed hippocampal patterns that discriminate reward context are associated with enhanced associative binding. *Journal of Experimental Psychology: General*, *142*(4), 1264–1276. http://doi.org/10.1037/a0033609

Xue, G., Dong, Q., Chen, C., Lu, Z., Mumford, J. A., & Poldrack, R. A. (2010). Greater Neural Pattern Similarity Across Repetitions Is Associated with Better Memory. *Science (New York, N.Y.)*, *330*(6000), 97–101. http://doi.org/10.1126/science.1193125

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., & Stark, C. E. L. (2010). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*, n/a–n/a. http://doi.org/10.1002/hipo.20808

Yassa, M. A., Mattfeld, A. T., Stark, S. M., & Stark, C. E. L. (2011). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences*, *108*(21), 8873–8878. http://doi.org/10.1073/pnas.1101567108

Yassa, M. A., & Stark, C. (2009). A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe☆. *NeuroImage*, *44*(2), 319–327. http://doi.org/10.1016/j.neuroimage.2008.09.016

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525. http://doi.org/10.1016/j.tins.2011.06.006

Yushkevich, P. A., Wang, H., Pluta, J., Das, S. R., Craige, C., Avants, B. B., … Mueller, S. (2010). Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI. *NeuroImage*, *53*(4), 1208–1224. http://doi.org/10.1016/j.neuroimage.2010.06.040