Title

A functional, behavioral, and model-based investigation of human visual memory

Permalink

https://escholarship.org/uc/item/1m78v97s

Author

Huffman, Derek

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

A functional, behavioral, and model-based investigation of human visual memory

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Derek James Huffman

Dissertation Committee:
Professor Craig E.L. Stark, Chair
Associate Professor Norbert J. Fortin
Associate Professor John F. Guzowski

2016

DEDICATION

To my wife Jessica

and my daughter June

you two are the center of my universe.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First and foremost I would like to thank my advisor, Dr. Craig Stark. Thank you for training me to be a critical and reflective scientist. You have taught me a great deal about how to approach a dataset and how to approach science in general. I appreciated your patience and I could not have asked for a better mentor. I am tremendously proud to say that I was fortunate enough to receive my Ph.D. training under your guidance.

I would like to acknowledge other members of the laboratory as their feedback has been instrumental to the research and ideas presented in this dissertation. I thank Shauna Stark for all of your feedback and advice along the way. I thank the graduate students that came before me for setting a high standard to which I aspired. I thank Veronique Boucquey for countless discussions about the brain, psychology, philosophy, and life—your support, feedback, and friendship have been invaluable. I am tremendously grateful for assistance with data collection from Patricia Place and Samantha Rutledge. I also thank Jessica German for assistance in data collection, data analysis, and for discussions about the behavioral experiments in Chapter 4.

I thank my committee for their valuable feedback on the experiments in this dissertation. I am particularly grateful to Drs. Norbert Fortin and John Guzowski for serving on my committee the entire time I have been at UCI and for helpful discussions at meetings and at journal club. I thank Drs. Steve Small and Alyssa Brewer for helpful comments at my advancement exam. I thank past and present members of the Medial Temporal Lobe Journal Club (aka Hebb Club) for invaluable discussions about the neurobiology of learning and memory. I thank Dr. Tim Allen for countless conversations about specific projects and about the field of neuroscience in general. I thank the Center for the Neurobiology of Learning and Memory for providing a rich scientific culture in which to study the neural correlates of memory. I thank Dr. Mike Yassa and members of his laboratory, especially Zach Reagh, for always sharing ideas with our laboratory. I thank Drs. Larry Cahill and Norm Weinberger for teaching me the importance of history and of the awareness of my assumptions.

I thank Tyler Sutterley for his friendship and support over the years. I thank Jeff and Nan Collett for introducing me to the CNLM. I thank my past basketball and rowing coaches—Chris Madigan, Pat Gleason, Larry Moore, Zach Johnson—and teammates for teaching me the importance of commitment and teamwork. I thank my entire family for always being there for me. I thank my parents for teaching me that I could achieve my goals through hard work and perseverance. I thank my brothers, Roger and Brian, for all that they have taught me over the years. Finally, I thank my wife Jessica and my daughter June for their unwavering support and for teaching me the true meaning of life.

CURRICULUM VITAE

Derek James Huffman

EDUCATION

Doctor of Philosophy in Neurobiology and Behavior   **2016**
University of California, Irvine

Bachelor of Science in Psychology   **2008**
University of California, San Diego

RESEARCH EXPERIENCE

Graduate Student Researcher   **2010–2016**
University of California, Irvine (Advisor: Craig Stark, Ph.D.)

Research Assistant   **2009–2010**
University of Southern California (Advisor: Mara Mather, Ph.D.)

Undergraduate Research Assistant   **2008**
University of California, San Diego (Advisor: Adam Aron, Ph.D.)

AWARDS AND NOTABLE ACHIEVEMENTS

Varsity Crew Captain (2005–2006). Orange Coast College.

Varsity Student Athlete: Crew (2006–2008). University of California, San Diego.

John W. Haycock Memorial Graduate Student Travel Award (2013). Center for the Neurobiology of Learning and Memory, University of California, Irvine.

Ralph Waldo Gerard Prize for Excellence in the History of Neuroscience (2014). University of California, Irvine.

Best Poster Award (2014). 19th Annual Meeting of the International Society for the History of the Neurosciences, Brussels, Belgium.

MEMBERSHIPS

International Society for the History of the Neurosciences
Society for Neuroscience

PEER-REVIEWED PUBLICATIONS

Nashiro, K., Sakaki, M., **Huffman, D.**, and Mather, M. (2013). Both younger and older adults have difficulty updating emotional memories. *The Journal of Gerontology, Series B: Psychological Sciences and Social Sciences*, 68(2):224–227.

Huffman, D. J. and 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.

Bennett, I. J., **Huffman, D. J.**, and Stark, C. E. L. (2015). Limbic tract integrity contributes to pattern separation performance across the lifespan. *Cerebral Cortex*, 25(5):2988–2999.

PUBLISHED ABSTRACTS AND CONFERENCE PRESENTATIONS

Huffman, D. J. and Stark, C. E. L. (October 2012). Multivariate investigation of domain specificity in human medial temporal lobe activity. Poster presented at the Society for Neuroscience Annual Meeting, New Orleans, LA.

Huffman, D. J. and Stark, C. E. L. (December 2012). Multivariate investigation of category representation in the human temporal lobe. Talk given at the Center for the Neurobiology of Learning and Memory Annual Conference, University of California, Irvine.

Huffman, D. J. and Stark, C. E. L. (January 2013). Fornix integrity predicts behavioral pattern separation performance. Talk given by Craig Stark at the Park City Annual Winter Conference on the Neurobiology of Learning and Memory, Park City, UT.

Huffman, D. J. and Stark, C. E. L. (January 2013). Multivariate investigation of category representation in the human temporal lobe. Talk given at the Park City Annual Winter Conference on the Neurobiology of Learning and Memory, Park City, UT.

Stark, C. E., **Huffman, D. J.**, Stark, S. M., and Yassa, M. A. (February 2013). Medial temporal lobe cortical thickness measurement using diffeomorphic registration in aging and mild cognitive impairment. Poster presented at the Annual Meeting of International Neuropsychological Society, Hawaii.

Bennett, I. J., **Huffman, D. J.**, and Stark, C. E. L. (November 2013). Integrity of limbic tracts contributes to episodic memory performance across the lifespan. Poster presented at the Society for Neuroscience Annual Meeting, San Diego, CA.

Huffman, D. J. and Stark, C. E. L. (November 2013). Multivariate pattern analysis of category representation in the human medial temporal lobe. Poster presented at the Society for Neuroscience Annual Meeting, San Diego, CA.

Huffman, D. J. (July 2014). Localization versus integration approaches to human neu-

roimaging since 1945. Poster presented at the Annual Meeting of the International Society for the History of the Neurosciences, Brussels, Belgium.

Boucquey, V. K., Allen, T. A., **Huffman, D. J.**, Fortin, N. J., and Stark, C. E. L. (November 2014). Memory for sequences of events shows bilateral hippocampal and medial prefrontal cortex activity in humans. Poster presented at the Society for Neuroscience Annual Meeting, Washington, DC.

Huffman, D. J. and Stark, C. E. L. (April 2015). Evidence for the representation of context in human parahippocampal cortex. Talk given at the Center for the Neurobiology of Learning and Memory Annual Conference, University of California, Irvine.

Allen, T. A., Boucquey, V. K., Quirk, C., **Huffman, D. J.**, Fortin, N. J., and Stark, C. E. L. (January 2015). A hippocampal-prefrontal system underlies memory for sequences of events in rats and humans. Talk given by Tim Allen at the Park City Annual Winter Conference on the Neurobiology of Learning and Memory, Park City, UT.

Boucquey, V. K., Allen, T. A., **Huffman, D. J.**, Fortin, N. J., and Stark, C. E. L. (October 2015). Hippocampus and medial prefrontal cortex show activity and functional connectivity during memory for sequences of events. Poster presented at the Society for Neuroscience Annual Meeting, Chicago, IL.

Huffman, D. J. and Stark, C. E. L. (October 2015). Evidence for the representation of context in human parahippocampal cortex and retrosplenial cortex. Poster presented at the Society for Neuroscience Annual Meeting, Chicago, IL.

INVITED LECTURES AND PRESENTATIONS

Computational models in cognitive neuroscience: tools for studying multiple memory systems
Neuroblitz Seminar Series, University of California, Irvine

April 2011

Multivariate investigation of category representation in the human medial temporal lobe
Neuroblitz Seminar Series, University of California, Irvine

May 2012

Multivariate investigation of category representation in the human medial temporal lobe
Brain Mapping Colloquium Series, University of California, Irvine

October 2012

Multivariate investigation of category representation in the human temporal lobe
Neuroblitz Seminar Series, University of California, Irvine

January 2013

Multivariate pattern analysis of category representation   June 2013
in the human medial temporal lobe
Ceremony for the John Haycock Memorial Graduate Student Travel Award, University
of California, Irvine

A brief history of analytic approaches to human neu-   January 2014
roimaging: from activation to connectivity to multivari-
ate pattern analysis
Ceremony for the Ralph Waldo Gerard Prize for Excellence in the History of Neuro-
science, University of California, Irvine

Multivariate pattern analysis of category representation   January 2014
in the human medial temporal lobe
Neuroblitz Seminar Series, University of California, Irvine

Investigation of representation in the human medial   April 2014
temporal lobe
Associated Graduate Student Symposium, University of California, Irvine

Investigation of the effects of test format on perfor-   December 2014
mance of the Mnemonic Similarity Task
Neuroblitz Seminar Series, University of California, Irvine

Investigation of representation in human parahip-   October 2015
pocampal cortex and retrosplenial cortex
Center for Neuroscience, University of California, Davis

A functional, behavioral, and model-based investigation   May 2016
of human visual memory
Doctoral Thesis Defense, University of California, Irvine

ACADEMIC SERVICE

Event Volunteer: A Memorable Evening   2011
Center for the Neurobiology of Learning and Memory, UC Irvine

Co-Coordinator of Neuroblitz Seminar Series   2012–2013
Department of Neurobiology and Behavior, UC Irvine

Event Volunteer: Evenings to Remember   2013
Center for the Neurobiology of Learning and Memory, UC Irvine

Coordinator of the Medial Temporal Lobe Journal Club   2013–2014
Center for the Neurobiology of Learning and Memory, UC Irvine

School Tour Docent   2012–2015
Center for the Neurobiology of Learning and Memory, UC Irvine

TEACHING EXPERIENCE

Graduate Student Instructor   **Spring 2010, Winter 2011**
Bio Sci N113L: Neurobiology and Behavior Laboratory   University of California, Irvine

Laboratory Leader   **2012–2014 (5 Quarters)**
Bio Sci N113L: Neurobiology and Behavior Laboratory   University of California, Irvine

Laboratory Developer and Writer   **2012–2014**
Bio Sci N113L: Neurobiology and Behavior Laboratory   University of California, Irvine

Laboratory Manual Editor   **2013–2014**
Bio Sci N113L: Neurobiology and Behavior Laboratory   University of California, Irvine

Mentor to Undergraduate Research Assistant   **2014–2015**
Mentee: Jessica German   University of California, Irvine

Invited Lecturer   **Winter 2015**
Bio Sci N255: History of Neuroscience (graduate course)   University of California, Irvine

Invited Lecturer   **Winter 2014, Winter 2016**
Bio Sci N119: History of Neuroscience   University of California, Irvine

COMPUTER PROGRAMMING

Python, R, GNU Octave, Psychtoolbox, Typesetting: LaTeX
Proficient on Debian-based GNU/Linux operating systems

NEUROIMAGING

Certified Philips 3T MRI Operator, University of California, Irvine
Proficient user of AFNI and PyMVPA

ABSTRACT OF THE DISSERTATION

A functional, behavioral, and model-based investigation of human visual memory

By

Derek James Huffman

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2016

Professor Craig E.L. Stark, Chair

Our ability to remember the events of our lives relies upon the formation of associations among the "what", "where", and "when" components of the event. Decades of research have focused on elucidating the contributions of subregions of the medial temporal lobe to memory for events. The results of this work converged on the notion that it would be fruitful to investigate the representation of information within subregions of the medial temporal lobe. Specifically, there is consensus that the hippocampus sits at the apex of a cortical circuit, which gives it unparalleled access to all aspects of event processing. In contrast, subregions of the adjacent medial temporal lobe cortex—namely, perirhinal and parahippocampal cortex— are hypothesized to be involved in the representation of object ("what") and contextual ("where") aspects of events, respectively. Recent studies have suggested that retrosplenial cortex, a region that has been largely absent from memory models, is also necessary for memory for events and that it is functionally related to parahippocampal cortex.

In four functional magnetic resonance imaging experiments, we investigated the hypothesis that parahippocampal cortex and retrosplenial cortex are preferentially involved in the representation of contextual information and that perirhinal cortex is preferentially involved in the representation of object information. Overall, the results of our experiments support our hypothesis; however, our results are incompatible with a simple functional dissociation

between these regions. Furthermore, our results provide evidence for the influence of low-level stimulus features in the representation of contexual information, suggesting that future research should aim to further investigate invariant context representation.

In a behavioral and model-based experiment, we investigated how memory changes as a result of healthy aging. Previous research has suggested that healthy aging is accompanied by an impaired ability to form stimulus-stimulus associations (e.g., "what-where") with a relative sparing of familiarity for the "what" component of events. An alternative, but not mutually exclusive, hypothesis is that healthy aging is accompanied by an impaired ability to encode stimulus features. Our results support the latter hypothesis, thus providing an alternative framework for the investigation of cognitive aging.

INTRODUCTION

Decades of research have converged on the notion that the medial temporal lobes are critically important for declarative memory—e.g., memory for the events of our lives—and that there is at least some degree of functional specialization between subregions of the medial temporal lobe. Specifically, many theories suggest that parahippocampal cortex is involved in memory for and the representation of contextual and spatial aspects of events (e.g., "where"), perirhinal cortex is involved in memory for and in the representation of items (e.g., "what"), and the hippocampus is involved in memory for all aspects of an event (e.g., "what-where-when"). Additionally, beyond the medial temporal lobe, recent theories are beginning to suggest that retrosplenial cortex, a subregion of posterior cingulate cortex, is involved in memory for and in the representation of context and spatial information. We performed four functional magnetic resonance imaging studies to investigate the information contained within patterns of activity in the medial temporal lobe and retrosplenial cortex.

Aim 1: Investigate whether parahippocampal cortex and retrosplenial cortex are preferentially involved in processing scene information. In two functional magnetic resonance imaging experiments, we investigated representations in response to different categories of images. In a between-subject design, participants viewed pictures of faces and objects or faces and scenes. We used multivariate pattern analysis to show that patterns of activity in perirhinal cortex, parahippocampal cortex, and retrosplenial cortex could be used to classify individual trials on which participants viewed faces versus objects and faces versus scenes. Consistent with our predictions, both parahippocampal cortex and retrosplenial cortex exhibited significantly better classification performance for faces versus scenes than for faces versus objects. Additionally, our results suggest that patterns of activity in parahippocampal cortex and retrosplenial cortex contain similar information on a trial-by-trial basis in our experiment with faces and scenes but not our experiment with faces and objects. These results are consistent with a role for parahippocampal cortex and retrosplenial

1

cortex in scene representation. Furthermore, these results suggest that these regions share information in a stimulus-dependent manner. Critically, however, the finding of significant classification accuracy in parahippocampal cortex and retrosplenial cortex for faces versus objects suggests that these regions are not tuned solely for context or scene processing.

Aim 2: Investigate the representation of individual items and contexts in the medial temporal lobe and retrosplenial cortex. Additionally, investigate whether representations are stable across different versions of the same items and contexts. The context-guided object association task (e.g., McKenzie et al., 2014) has provided a useful task to investigate the representation of events within subregions of the rodent medial temporal lobe. We created two human versions of the context-guided object association task for use in functional magnetic resonance imaging experiments. In our first experiment, we used distinct items and contexts similar to the rodent task. We provide evidence for the representation of context in the medial temporal lobe and retrosplenial cortex; however, we found similar results in primary visual cortex, suggesting that our results could be influenced by low-level visual features. In our second experiment, we tested whether representations in the medial temporal lobe and retrosplenial cortex would exhibit invariance across different versions of the same items and contexts. Our results provide little evidence for invariant context representation in the medial temporal lobe and retrosplenial cortex, suggesting that the results of our first experiment were highly-dependent upon low-level visual differences between stimuli. In contrast, our results provide novel evidence for invariant object representation in perirhinal cortex. The results of our experiments should constrain the experimental designs of future studies that aim to investigate the representation of context.

Aim 3: Investigate performance on a forced-choice version of a Mnemonic Similarity Task in healthy younger adults and the changes that occur as a result of healthy aging. Additionally, use computational models to provide a potential mechanistic account of the empirical results. Previous research from our laboratory

has shown that healthy older adults are impaired on tasks that require discriminating between previously viewed images and similar lure images. We extended previous results from the old/new and the old/similar/new test formats to the forced-choice format. Additionally, we used a class of models from mathematical psychology—global matching models—to show that impaired encoding provides a possible mechanistic explanation for the observed impairments in healthy older adults. Altogether, these results advance our understanding of the memory changes associated with healthy aging and the modeling results provide an alternative framework for interpreting performance on the Mnemonic Similarity Task.

Chapter 1

Background and Significance

Our memories provide us with a sense of personal history, thus governing the connections that we feel with our loved ones. Accordingly, our memories are at the core of who we are. Sadly, there are neurodegenerative diseases that rob individuals of their memories. The loss of one's personal history is devastating not only for patients but also for their loved ones. Therefore, the study of human memory is relevant both in healthy individuals and in patients with memory impairments. If we can increase our understanding of how memory works at a behavioral, computational, or implementational (i.e., brain) level, then we will move one step closer to understanding how to fix memory in patients with memory impairments. The goal of this dissertation was to investigate human visual memory and perception using functional neuroimaging, behavioral experiments, and computational modeling in an effort to further our understanding of memory in healthy younger adults and the changes to memory that occur as a result of healthy aging. In this chapter, I provide relevant background material before discussing the experiments in this dissertation in subsequent chapters.

1.1 An Introduction to Memory Systems

1.1.1 Declarative versus Nondeclarative Memory

In 1957, Scoville and Milner published a paper that is arguably the most influential paper ever published on the neurobiology of learning and memory. William Scoville removed portions of the hippocampus, amygdala, and adjacent medial temporal lobe (MTL) cortex in several patients with intractable epilepsy, including the famous patient, H.M. They reported devastating memory impairments following MTL damage. Subsequent work by Milner revealed that H.M. was only impaired on certain types of memory tasks, thus providing neurobiological evidence that memory is not a single entity (Milner, 1962; Milner et al., 1968). For example, his overall intelligence remained intact, and he performed as well as controls on a mirror-tracing task, in which participants were instructed to draw between two lines while looking in a mirror (Milner, 1962). The number of errors (an error was recorded when participant drew outside the lines) provided an index of memory performance. Participants performed the same experiment across multiple days, thus allowing testing both of learning rate and memory retention. Milner found that H.M. performed as well as controls on the task, including across-session retention; however, he had no memory that he had previously performed the task. These findings provided evidence for a single dissociation—that is, impaired performance on one type of task with spared performance on another type of task following localized lesions—thus providing neurobiological evidence for the hypothesis of multiple memory systems. Following Milner's reports (Milner, 1962; Milner et al., 1968), researchers focused on further characterizing memory systems.

Cohen and Squire (1980) aimed to further uncover spared memory ability in patients with amnesia. They had participants perform a task in which they read words in a mirror. With training, participants increased the rate at which they could read the words. Interestingly, patients with amnesia showed a similar rate of learning, and they retained their memories

for at least three months. In contrast, they found that patients with amnesia were severely impaired at learning word pairs. Taken together, Cohen and Squire (1980) argued that there were at least two forms of memory: 1) "declarative" or memory for "data-based material", 2) "procedural" memory. They also posited that their results extended previous research by showing that the spared memory function was not limited to motor tasks (e.g., mirror-drawing task; Milner, 1962; Milner et al., 1968).

Packard and McGaugh (1996) used temporary inactivation techniques in rodents to provide strong neurobiological evidence for the existence of multiple memory systems. They utilized a cross-maze in which they closed one of the arms of the maze, effectively creating a modified T-maze (the "testing" arm was closed during training, and the "training" arm was closed during testing; see Figure 1.1). They trained rats that one of the arms of the maze contained a reward. There were two strategies that the rats could use to learn the task: 1) "place" (e.g., the East arm is rewarded), 2) "response" (e.g., turn right). During the test phase, the rats started in the arm opposite the training arm, thus allowing testing of the learning strategy used by the rats.

After 8 days of training, the majority of saline-infused control rats used the "place" strategy. Inactivation of the hippocampus resulted in chance performance (half of the rats turned left and half of the rats turned right); whereas inactivation of the caudate resulted in behavior that was indistinguishable from the saline-infused controls. After 16 days of training, the majority of saline-infused control rats used the "response" strategy, thus suggesting a strategy shift with increased training. Hippocampal inactivation resulted in
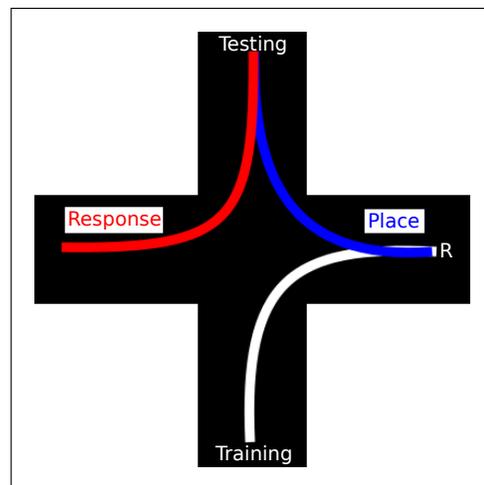
Figure 1.1: Diagram of the cross-maze paradigm (R=Reward).

performance that was indistinguishable from saline infused controls; however, the majority of the caudate-inactivated animals used the "place" strategy. These results suggest that after 8 days of training the rats use a hippocampal-dependent "place" strategy, and that after 16 days of training the rats use a caudate-dependent "response" strategy. The behavior of the caudate-inactivated animals (after 16 days of training) suggests that the hippocampal-dependent "place" strategy is still available but that it is in competition with the caudate-dependent "response" strategy. These findings provide evidence for a double dissociation between the hippocampus and the caudate, thus providing stronger evidence for dissociable memory systems.

Taken together, researchers leveraged the findings from studies of animals and patients with amnesia to suggest that there is a dissociation between declarative (or relational) memory and nondeclarative memory (e.g., Cohen and Squire, 1980; Squire, 1992; Squire and Zola, 1996; Eichenbaum, 2000). These examples suggest that the MTL is critical for declarative memory but not for nondeclarative memory. There are several subtypes of nondeclarative memory, but detailed discussion of nondeclarative memory is beyond the scope of this chapter. Sherry and Schacter (1987) suggested that the brain evolved multiple memory systems that operate to serve "functionally incompatible" goals. They postulated that nondeclarative and declarative memory systems—in their words, "System I" and "System II", respectively—are two such systems. For example, they argued that System I would benefit from "invariance" over repeated experiences, whereas System II would benefit from maintaining "variance" over repeated experiences. Invariance would allow an organism to extract statistical regularities of the world while variance would allow an organism to form unique memories for individual events. Given this opposition, they hypothesized that these systems rely on independent brain regions. Sherry and Schacter (1987) thus provide one possible explanation for the selective declarative memory impairment following MTL damage. Subsequent research aimed to further parcel the unique contributions of subregions of the MTL—namely, the MTL cortex and the hippocampus—to declarative memory. Before discussing these studies, I will

provide a brief background of the anatomy of the MTL.

1.1.2   Anatomy of the Medial Temporal Lobe

The MTL is composed of several subregions: parahippocampal cortex (PHC; called postrhinal cortex in rodents), perirhinal cortex (PRC), entorhinal cortex (EC), and the hippocampal region. The hippocampal region is made up of the dentate gyrus (DG), cornu ammonis fields (CA1–4; research has focused on CA1 and CA3), and the subiculum. The EC is the major input and output structure of the hippocampus, and it receives projections from PHC and PRC. Research has suggested that PRC and PHC receive the majority of their anatomical projections from the "ventral visual stream" and the "dorsal visual stream," respectively (see Figure 1.2; Suzuki and Amaral, 1994; Ungerleider, 1995; Squire and Zola, 1996; Buffalo et al., 2006; Diana et al., 2007; Eichenbaum et al., 2007; Ranganath and Ritchey, 2012).

The ventral visual stream is implicated in playing a role in "item" or "what" processing. It has been suggested that the primary visual cortex represents relatively simple features (i.e., lines), and that the information coded along the ventral visual stream becomes progressively more conjunctive (Ungerleider, 1995; Cowell et al., 2010b). The inferior temporal cortex (IT) is hypothesized to be involved in the representation of objects. PRC receives the majority of its inputs from IT and other regions in the anterior temporal cortex (Suzuki and Amaral, 1994; Eichenbaum et al., 2007; Ranganath and Ritchey, 2012). The dorsal visual stream is implicated in playing a role in "where" (Ungerleider, 1995) or "how" (Goodale and Milner, 1992) processing. PHC receives strong projections from posterior cingulate cortex (PCC) and retrosplenial cortex (RSC; Suzuki and Amaral, 1994; Buffalo et al., 2006). RSC sits at the interface between the "dorsal visual stream" and the MTL. Interestingly, RSC and PCC have only very weak projections to PRC (Suzuki and Amaral, 1994). Taken together, these results have led recent theoretical accounts to suggest that these regions differ in terms
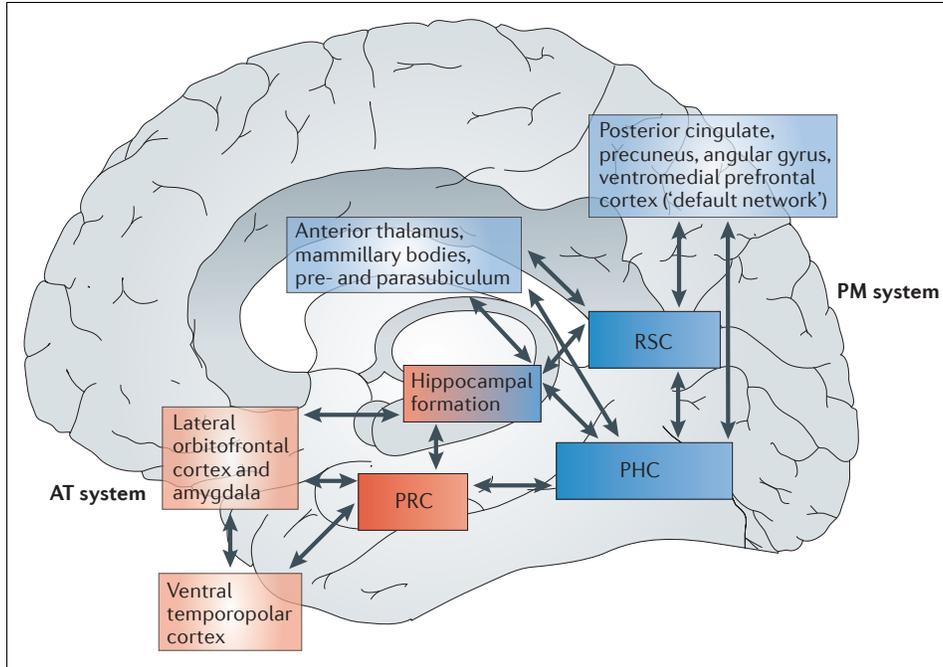
Figure 1.2: PHC is part of a broader posterior medial (PM; represented in blue) system involved in the memory of context while PRC is part of a broader anterior temporal (AT; represented in red) system involved in memory of items. RSC=retrosplenial cortex, PHC=parahippocampal cortex, PRC=perirhinal cortex. Figure from (Ranganath and Ritchey, 2012).

of the information that they represent—namely, PRC is hypothesized to be preferentially tuned for "item" (or object) memory while PHC is hypothesized to be preferentially tuned for context or scene memory (e.g., Diana et al., 2007; Eichenbaum et al., 2007; Ranganath and Ritchey, 2012).

PRC and PHC differ in terms of their connectivity with EC. Specifically, anatomical studies in the rodent and monkey have shown that PRC has stronger projections to lateral entorhinal cortex while PHC has stronger projections to medial entorhinal cortex. Accordingly, it has been suggested that lateral and medial EC are differentially involved in memory for items and spatial contexts, respectively (e.g., Knierim et al., 2006; Ranganath and Ritchey, 2012; Ritchey et al., 2015). However, recent theoretical accounts have suggested that the classic "what" and "where" dissociations between MTL cortical regions are too simplistic. Specifically, functional and anatomical investigations have shown interconnectivity between

MTL cortical regions (e.g., there are projections between PHC and PRC; Lavenex and Amaral, 2000; Buffalo et al., 2006) and MTL cortical regions have been shown to contain information about spatial (e.g., in perirhinal and lateral EC) and non-spatial (e.g., in medial EC and PHC) aspects of the task (e.g., Knierim et al., 2013; McKenzie et al., 2015; Keene et al., 2016). These are relatively recent views, however, so further discussion is beyond the scope of this chapter (for a more thorough discussion see **Chapter 3** and **Chapter 5**).

The hippocampus receives highly processed information from every sensory modality, which has led some researchers to suggest that the hippocampus is involved in memory because it is a site of the brain where information is combined into a unique representation (Squire, 1992; Ungerleider, 1995; O'Reilly and Munakata, 2000, Chapter 9; Diana et al., 2007; Eichenbaum et al., 2007; Ranganath and Ritchey, 2012). That is, the hippocampus is able to form multimodal conjunctive representations. In fact, some theories suggest that the hippocampus is specialized for the formation of conjunctive item-in-context representations, thus subserving a selective role in associative memory performance (Eichenbaum et al., 2007; Ranganath, 2010; Ranganath and Ritchey, 2012; Ritchey et al., 2015). An alternative hypothesis, also based on anatomical considerations, is that the hippocampus can reinstate cortical patterns of activity in distributed regions throughout the brain (Teyler and DiScenna, 1986; Teyler and Rudy, 2007).

The major anatomical projection from EC to the hippocampus is the perforant path, which projects from EC layer II to DG and CA3 (see Figure 1.3; Amaral, 1993; Wilson et al., 2006). The DG is roughly 5 times larger than the EC, which leads to sparser representations in the DG via "expansion" (McNaughton and Morris, 1987). The DG projects to CA3 via the mossy fibers (Amaral, 1993; Wilson et al., 2006), which are often referred to as "detonator synapses" (McNaughton and Morris, 1987) because they are extremely strong. More than 95% of the synapses in CA3 are from the auto-associative fibers (i.e., self-projections; Wilson et al., 2006). Taken together, CA3 is uniquely situated to either form unique patterns of
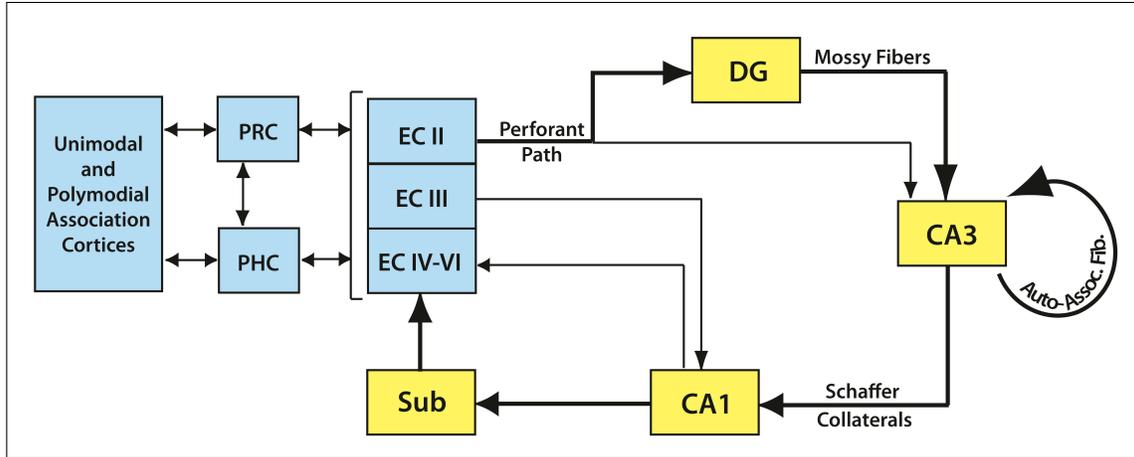
Figure 1.3: A simplified schematic of connectivity of the MTL (figure modified from Wilson et al., 2006; courtesy of J.W. Lacy). PRC=perirhinal cortex, PHC=parahippocampal cortex, ERC=entorhinal cortex, PP=perforant path, SC=Schaffer collaterals, DG=dentate gyrus.

activity—via mossy fiber input—or to complete partially degraded input patterns—via the auto-associative fibers (e.g., O'Reilly and McClelland, 1994). CA1 receives input from both CA3—via the Schaffer collaterals—and EC layer III (Amaral, 1993; Wilson et al., 2006). Theoretical accounts have suggested that these anatomical properties allow CA1 to act a as a comparator of the patterns of activity between EC and CA3 (Hasselmo et al., 1995). The subiculum is thought to be a major output structure—from CA1 to EC layer IV-VI—of the hippocampus (Amaral, 1993; Wilson et al., 2006). The function of subregions of the hippocampus will be further explored later.

Recent theories have suggested that we need to expand our view of the declarative memory system to include regions outside of the MTL (e.g., Vann et al., 2009; Aggleton, 2010; Ranganath and Ritchey, 2012; Bucci and Robinson, 2014; Ritchey et al., 2015). For example, as mentioned above, RSC is anatomically situated to be involved in context processing (based on its connectivity with the dorsal processing stream) and in declarative memory (based on its connectivity with the MTL). RSC is reciprocally connected to the hippocampus, PHC, and anterior thalamic nuclei, which are critically important for declarative memory (Vann et al., 2009; Aggleton, 2010). Indeed, lesions to RSC have been shown to result

in "retrosplenial amnesia" (Valenstein et al., 1987) and RSC atrophy as been observed in patients with Alzheimer's disease (Buckner et al., 2005; Fennema-Notestine et al., 2009; Tan et al., 2013). Moreover, RSC hypometabolism has been observed in the early stages of Alzheimer's disease (Minoshima et al., 1997; Buckner et al., 2005; Villain et al., 2008) and there is evidence to suggest that the degree of memory decline is associated with the degree of RSC hypometabolism (Desgranges et al., 2002; Buckner et al., 2005). Furthermore, a recent model suggests that RSC is critical for the formation of stimulus-stimulus associations, a function traditionally ascribed to the hippocampus (Bucci and Robinson, 2014). Thus, a more complete understanding of the contribution of structures beyond the MTL, especially RSC, to declarative memory is of keen interest. I will later expand on this argument (see **Chapter 2**, **Chapter 3**, **Chapter 5**), but for now I will discuss theories that have attempted to understand the differential involvement of the hippocampus and MTL cortex in declarative memory, beginning with studies of recollection and familiarity which dominated the field in the 1990s and the early 2000s.

1.1.3   Recollection and Familiarity

Recognition memory tasks have received much attention over the past 40 years, and they have recently been used to suggest that the hippocampus and PRC play complementary roles in memory formation. Recognition memory tasks test participants' ability to discriminate between stimuli that they have and have not previously encountered. For example, participants might study a list of words followed by a test on which they indicate whether each word is "old" or "new." Dual-process models of recognition memory (Mandler, 1980; Yonelinas, 1994, 1997, 2002) posit that recognition memory judgments can be made using two processes: recollection and familiarity. Recollection is defined as the process of recalling details about the study event. For example, when a participant studies the item "apple," they might imagine visiting an apple orchard. When tested on the item "apple," the par-

ticipant might recall that they imagined visiting an apple orchard. Thus the participant would respond "old," based on recollection. On the other hand, the participant might have a strong sense of familiarity for the word "table," but not be able to recall anything about studying the word. The Yonelinas model (Yonelinas, 1994, 1997, 2002) posits that whenever a participant recalls something about the studied item, they make their decision based solely on recall; whereas, if they cannot recall anything about the studied item, then they make their decision based solely on familiarity.

Following the development of dual-process theories of recognition memory, researchers have used lesion techniques, electrophysiological recordings, human neuroimaging, and studies of patients with selective brain damage in an attempt to provide evidence for functional differences between PRC and the hippocampus. A prominent theory that emerged from the efforts of these studies suggested that the PRC is specialized for familiarity while the hippocampus is specialized for recollection (for reviews see: Aggleton and Brown, 1999; Brown and Aggleton, 2001; Yonelinas, 2002; Diana et al., 2007; Eichenbaum et al., 2007; Ranganath, 2010; Yonelinas et al., 2010; Brown et al., 2010; Ranganath and Ritchey, 2012). These theories also note a role for the hippocampus in the formation of associative memories (e.g., "what-what", "what-where", "what-where-when"), while perirhinal cortex is hypothesized to play a larger role in item memory (e.g., "what"). However, other studies have provided evidence to suggest that the hippocampus is involved in recollection, familiarity, and associative and single item memory (e.g., Stark and Squire, 2003; Bayley et al., 2008; Kirwan et al., 2008; Jeneson et al., 2010; Wais et al., 2010; Wixted et al., 2010). Accordingly, other theories suggest that the hippocampus and MTL cortical regions are broadly involved in declarative memory, including a role for the hippocampus in familiarity (e.g., Squire et al., 2007; Wixted et al., 2010; Wixted and Squire, 2011).

The issue of recollection and familiarity differences between the hippocampus and PRC has been the source of a fierce debate in the literature; however, there are points of agreement

between competing theories. For example, proponents of both theories have suggested that anatomical connections in the hippocampal-cortical network may give rise to differences in the information processed in different subregions of the MTL. For example, both the Binding of Item in Context model (Eichenbaum et al., 2007; Diana et al., 2007) and the Attributes of Memory model (Wixted and Squire, 2011), suggest that PHC may be relatively specialized for memory for context, spatial, or scene information, PRC may be relatively specialized for memory for item or object information (however, some models suggest that PRC is involved in processing of both scene and object information; Buffalo et al., 2006; Wixted and Squire, 2011), and the hippocampus is important for memory across domains. Taken together, both of these models suggest that it would be fruitful to investigate the representation of information within the MTL. The consensus of these views inspired us to investigate representations in the MTL (see **Chapter 2** and **Chapter 3**). As I will discuss in subsequent sections, computational theories provide a mechanistic account of differences between regions, which may provide clearer explanation than reliance on psychological theory (Rumelhart et al., 1986; O'Reilly and Munakata, 2000; O'Reilly and Norman, 2002; Norman and O'Reilly, 2003; Norman, 2010; Cowell et al., 2012). For example, it has been argued that the terms "recollection" and "familiarity" should be replaced by terms that precisely describe mechanistic—rather than psychological—processes (Norman, 2010).

1.2 An Introduction to Computational Theory

1.2.1 Global Matching Models

Researchers in the field of mathematical psychology have used a combination of modeling and behavioral experiments to investigate the nature of memory representations. For example, global matching models have been used extensively to investigate recognition memory (e.g.,

Hintzman, 1984, 1988, 2001; Murdock, 1982, 1995; for review see: Clark and Gronlund, 1996). Global matching models assume that participants calculate a global match between a test item and the stored contents of memory. If the global match is strong enough, then the participant responds "old," otherwise they respond "new." These models are referred to as "familiarity-only" models because they are single-process models that make decisions based on match of a test item to the contents of memory. Accordingly, global matching models are related to the signal detection theory framework (cf. Clark and Gronlund, 1996). However, rather than assuming differences in memory strength per se, global matching models assume that the global match (i.e., a memory strength signal) is determined by the similarity of a test item to the contents of memory (cf. Kahana, 2012, Chapter 3). There are many versions of global matching models, however the main difference between the models is whether they assume that memories are stored as independent traces (i.e., multiple-trace or exemplar-based models) or as a distributed trace (i.e., a single, composite memory or a prototype). We investigated whether global matching models could account for the performance of younger adults and healthy older adults on an item recognition memory test with targets and similar lures (see **Chapter 4**). Global matching models are abstract, meaning that they attempt to model latent cognitive representations without taking into account the neural machinery behind such computations. Other researchers have advocated for using knowledge of brain connectivity and function to guide development of computational models. In the next two sections, I will discuss biologically-inspired computational theories of the brain.

1.2.2 Psychological Similarity

Researchers have been studying "psychological similarity" for over 50 years (Rothkopf, 1957; Shepard and Chipman, 1970; Shepard, 1987; Cutzu and Edelman, 1998; Edelman, 1998, 2012). Psychological similarity is a broad concept that encompasses perceptual, conceptual, and semantic similarity. For example, Rothkopf (1957) investigated the effects of psycholog-

ical similarity on paired-associate learning. Participants performed a task in which stimulus pairs were presented sequentially with a brief delay between them, and they were instructed to indicate whether or not the two stimuli were the same. The number of between-stimulus errors (e.g., "same" response when presented with A and B) provided an index of psychological similarity. Due to the complex relationship between physical similarity and psychological similarity, Rothkopf (1957) noted that investigation of the relationship between memory and stimulus similarity should utilize measures of psychological similarity rather than measures of physical similarity. He showed that there was a relationship between difficulty of paired associates learning and psychological similarity.

Shepard and Chipman (1970) developed a hypothetical solution for mapping between physical and psychological space. Cognitive psychologists use the term "space" to refer to the dimensions of stimulus representation. Stimuli can vary across several dimensions (e.g., size, color, shape), and the dimensions make up the axes of the space. Representation of a stimulus can be thought of as a point in multidimensional space, where similar stimuli occupy nearby regions of space. Shepard and Chipman (1970) suggested that the relationship between physical and psychological space should adhere to a "second order isomorphism." A second order isomorphism states that the similarity between two stimuli in psychological space should be related to the similarity of the stimuli in physical space. For example, the representation of a face in psychological space (or in brain space) need not actually look like a face—in fact, if it did, such a representation would be referred to as a "first order isomorphism." Instead, the representational similarity between two images of faces should be greater than between either face image and an image of a scene. In this framework, generalization is permitted via the similarity of the stimuli in psychological space. Behavioral methods such as those used by (Rothkopf, 1957) can be used to test the second order isomorphism. The framework of the second order isomorphism was subsequently applied to representations in neural-network models.

Theoretical approaches have suggested that there are several ways that neurons can respond to stimuli. Taken together, these theories refer to how neurons—and, more broadly, brain regions—represent information. One view of representations suggests that neurons use rate coding, which refers to the idea that neurons code information based on changes in firing rate (Dayan and Abbott, 2005). Localist views of representation suggest that localized units represent specific features of stimuli (from the example in the previous paragraph: size, color, shape; in this example, the mapping between the coordinates of physical space and "neural" space would be one-to-one; Rumelhart et al., 1986, Chapter 3). Other theorists suggested that the brain utilizes "coarse-coded" distributed representations (e.g., Hinton, 1981; Churchland, 1986; Rumelhart et al., 1986, Chapter 3), also referred to as "population coding" (Dayan and Abbott, 2005). The basic idea of coarse coding is that each unit is broadly tuned and thus participates in the representation of several features. I will use the term "distributed representations" as a shorthand of referring to coarse-coded distributed representations. Distributed representations exhibit several advantages over localist representations: they are more efficient, they naturally represent the similarity structure of their environments and are able to generalize knowledge to novel stimuli, they are more accurate, and they exhibit graceful degradation (Hinton, 1981; Rumelhart et al., 1986, Chapters 3, 4, 12; O'Reilly and Munakata, 2000)

Rumelhart and Todd (1993) suggested that neural-networks could capture many properties of semantic cognition. They trained neural-networks to learn several facts about different animals. For example, they taught the networks that a canary is a bird, that it can fly, that it has feathers. After training the networks on several categories of stimuli—trees, flowers, birds, and fish—they tested how well the networks could represent novel exemplars from these categories. For example, after training the networks that a new stimulus was a bird, the networks correctly predicted that the new bird could fly and that it had wings and feathers. Similarly, the networks correctly predicted that the new bird did not contain non-bird features (e.g., gills). This highlights one utility of distributed representations, namely

generalization via proximity in "psychological" space.

In summary, distributed representations give rise to emergent properties, such as generalization (Rumelhart et al., 1986, Chapter 3). Distributed representations are not necessarily at odds with specialization of function—that is, it is possible that the brain is composed of relatively specialized regions (e.g., object processing regions), and that within each of these specialized regions, neurons make use of distributed representations (Rumelhart et al., 1986, Chapter 3). While generalization and representation of similarity are beneficial for rapidly predicting the world around us, being able to uniquely represent past experiences is a defining feature of episodic memory (Tulving, 1985). In the next section, I will discuss a computational theory for how the brain decreases interference, potentially allowing unique memory traces to be formed.

1.2.3   Pattern Separation

Pattern separation is a computational process that decreases the similarity of overlapping input patterns (e.g., O'Reilly and McClelland, 1994; McClelland et al., 1995; Norman and O'Reilly, 2003; Norman, 2010). For example, if two binary input patterns share 60% of the same active units (i.e., 40% change in input between the two patterns), then pattern separation would transform the patterns to share fewer than 60% of same the active units (i.e., greater than 40% change in output; blue region in Figure 1.4), pattern
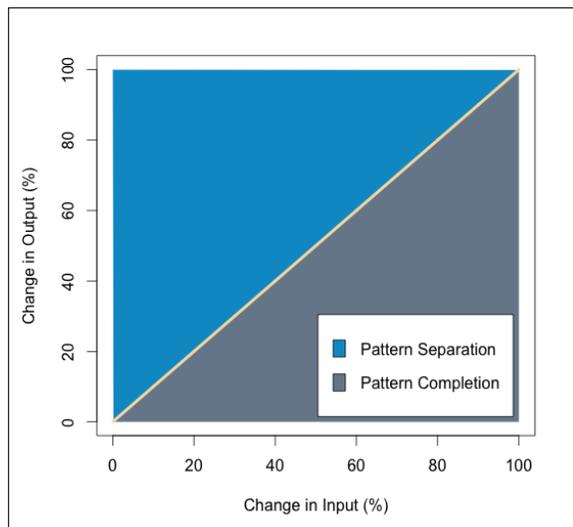
Figure 1.4: Pattern separation and pattern completion transform the similarity structure among input patterns.

18

completion would transform the patterns to share more than 60% of the same active units (i.e., less than 40% change in output; gray region in Figure 1.4), and the process of linearly tracking the change in input patterns would result in no transformation to the similarity between the input patterns (i.e., 40% change in output; tan line in Figure 1.4). Pattern separation decreases the similarity among input patterns and thus could serve as a computational mechanism for reducing similarity among related experiences. Pattern completion increases the similarity among input patterns, and could thus allow retrieval of a partially degraded input (Marr, 1971; McNaughton and Morris, 1987; O'Reilly and McClelland, 1994; Hasselmo et al., 1996). This section will focus on biologically-inspired computational modeling studies that have been used to advance the notion that the hippocampus is involved in pattern separation.

Before discussing computational models of the hippocampus, an example will be given to illustrate the computational trade-off between generalization and specificity. McCloskey and Cohen (1989) examined the performance of simple neural-networks—like those discussed in the previous section—on a memory interference task, the AB-AC list learning task. In this task, participants learn word pairs on two separate lists. Both lists consist of one of the same words (i.e., A) with a repaired second word (i.e., AB is learned on the first list, followed by AC on the second list). Behavioral studies in humans indicated that learning the AC list resulted in retroactive interference (i.e., worse performance when tested on the AB list); however, learning the AC list did not completely interfere with the memory of the AB pairs (Barnes and Underwood, 1959). In contrast, the neural-networks exhibited complete retroactive interference, a pattern of results that they referred to as "catastrophic interference." They argued—because the neural-network results did not match human behavior—that simple neural-networks were not appropriate models of the brain.

In contrast to the arguments advanced by McCloskey and Cohen (1989), McClelland et al. (1995) suggested that the combination of useful properties of neural-networks (e.g., the

ability to generalize to novel stimuli; Rumelhart and Todd, 1993) and the observation of catastrophic interference in a memory interference task suggests that the brain evolved complementary learning systems that operate under competing computational principles. For example, in an attempt to overcome catastrophic interference in neural-networks, French (1991) used sparser representations; however, while these networks exhibited lower levels of retroactive interference, they performed worse at generalizing and representing the similarity structure of their environments. McClelland et al. (1995) argued that generalization requires distributed representations and gradual learning, and that the formation of arbitrary associations requires the rapid formation of non-overlapping memories. Accordingly, McClelland et al. (1995) suggested that the brain makes use of complementary learning systems: 1) the neocortex, which utilizes a slow learning rate, 2) the hippocampus, which allows conjunctive, associative learning.

There are several properties of the hippocampus that make it functionally suited for pattern separation. First, physiological investigation suggested that the hippocampus has much sparser activation than the neocortex (e.g., Barnes et al., 1990). In particular, the DG exhibits very sparse activity ($\approx$ 0.5%; O'Reilly and McClelland, 1994; O'Reilly and Munakata, 2000, Chapter 9; Chawla et al., 2005). Mathematically, the lower the probability that a neuron is active, the lower the probability that it will be active in response to two stimuli (see Figure 1.5; O'Reilly and McClelland, 1994; O'Reilly and Munakata, 2000, Chapter 9; O'Reilly and Rudy, 2001). Second, DG contains
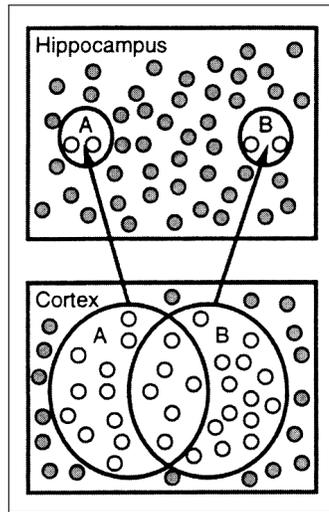
Figure 1.5: An illustration of distributed representations in the cortex and sparse-coded, pattern-separated representations in the hippocampus (figure from O'Reilly and Rudy, 2001). A and B represent patterns of activity in response to two stimuli.

roughly 5 times as many neurons as EC, which serves to decrease the similarity of overlapping input patterns through the process of "expansion" (McNaughton and Morris, 1987). Third, the mossy fibers sparsely project from DG to CA3 and they provide very strong input ("detonator synapses"; McNaughton and Morris, 1987), thus allowing further pattern separation (O'Reilly and McClelland, 1994). Fourth, the hippocampus (in contrast to the neocortex) uses a fast learning rate, which allows rapid formation of novel, pattern-separated representations (O'Reilly and Norman, 2002). Computational modeling studies of the hippocampus have shown that these properties serve to reduce the overlap of similar input patterns, thus serving as a potential mechanism for reducing interference between similar experiences (Treves and Rolls, 1992, 1994; O'Reilly and McClelland, 1994; Hasselmo et al., 1996; Hasselmo and Wyble, 1997; O'Reilly and Munakata, 2000, Chapter 9; Norman and O'Reilly, 2003; Norman, 2010).

Computational models have been used to suggest that network dynamics differ between encoding and retrieval modes (Treves and Rolls, 1992, 1994; O'Reilly and McClelland, 1994; Hasselmo et al., 1996; Hasselmo and Wyble, 1997). It has been suggested that DG is important for encoding but not for retrieval, due to its is bias toward forming unique patterns of activity even for similar stimuli. Additionally, the very strong mossy fiber synapses from DG to CA3 could drive CA3 into a pattern separation mode. In contrast, without DG input, CA3 can exhibit pattern completion tendencies due to its auto-associative connectivity. In fact, 95% of the synapses in CA3 are auto-associative connections (Wilson et al., 2006). Previous modeling studies have suggested that interactions between the hippocampus and the cholinergic system can serve to shift the hippocampus between encoding and retrieval modes (Hasselmo et al., 1996; Hasselmo and Wyble, 1997). For example, under high levels of acetlycholine, CA3 attractor dynamics are inhibited, thus biasing the hippocampal network toward pattern separation or encoding mode. In contrast, under low levels of acetlycholine, CA3 attractor dynamics are promoted, thus biasing CA3 into pattern completion or retrieval mode. Taken together, widespread interactions may play a role in shifting hippocampal pro-

cessing between pattern separation and pattern completion.

Computational models can provide mechanistic insight into functional differences between brain regions. For example, previous studies have investigated the computational differences between the MTL cortex and the hippocampus (Norman and O'Reilly, 2003; Norman, 2010). These investigations suggested that, due to its use of distributed representations, patterns of activity in the MTL cortex could be used to track the global similarity or summed similarity between a test item and studied items. Accordingly, their models of the MTL cortex are in fact an instantiation of a global matching model. In contrast, due its bias toward pattern separation, patterns of activity in the hippocampus were largely non-overlapping in response to studied and unstudied items; therefore, activity in the hippocampus cannot track the summed similarity between a test item and studied items. Norman (2010) argued that in testing the functional differences between the hippocampus and MTL cortex, the terms "recollection" and "familiarity" should be replaced with the predicted mechanistic differences between these regions. That is, future research should test the hypothesis that the MTL cortex can compute summed similarity while the hippocampus cannot because of differences in the degree of pattern separation in the MTL cortex and the hippocampus. One aim of the research in this dissertation was to provide empirical support for these mechanistic predictions (see **Chapter 2**).

1.2.4 Application of the Theory: Empirical Investigations of Hippocampal Pattern Separation

Lesion Studies

The lesion method has been used to assess the necessity of the hippocampus in performance of tasks that were designed to provide a behavioral assay of pattern separation. Taken to-

gether, these tasks will be referred to as "mnemonic similarity tasks" (cf. Stark et al., 2015). Pattern separation is a computational process—i.e., a process that decreases the overlap of input patterns—while discriminating between similar memories is a psychological process, which could arise from mechanistic processes other than pattern separation. Therefore, the studies discussed in this section will provide evidence for the necessity of the hippocampus in pattern separation only to the extent that discriminating between similar memories accurately assesses pattern separation. Lesion studies in rodents and human patients have provided converging evidence for a role of the hippocampus in performance on mnemonic similarity tasks (i.e., targets and similar lures) but not traditional item recognition memory tasks (i.e., targets and unrelated foils).

There is evidence that lesions of the rat dentate gyrus cause impaired discrimination between close—but not distant—locations in spatial memory tasks (Gilbert et al., 1998, 2001; Morris et al., 2012). Similarly, a recent study of a patient with selective hippocampal damage revealed impaired precision of spatial memory judgments on a virtual water maze task (Kolarik et al., 2016). Talpos et al. (2010) reported that hippocampal damage causes impaired mnemonic discrimination of very similar spatial patterns in a "trial unique, delayed nonmatching-to-location" task in rodents. Importantly, studies have shown that hippocampal damage does not impair perceptual discrimination of similar items, which has been used to suggest that the observed behavioral changes following damage to hippocampus are related to its role in declarative memory rather than the result of a simple perceptual deficit (Suzuki, 2009; but see Saksida and Bussey, 2010; Burke et al., 2012).

The complementary learning systems (CLS) model predicts that hippocampal damage would cause impaired on performance on "old"/"new" memory tests that include targets and similar lures while sparing performance on "old"/"new" memory tests that include targets and unrelated foils (Norman and O'Reilly, 2003). This occurs because the neocortex (e.g., PRC) assigns similar representations to similar items, which results in a greater probability of

incorrectly responding "old" to related lures (compared to the hippocampus). As predicted by the CLS model, Holdstock et al. (2002) showed that a patient with selective hippocampal damage was impaired on an "old"/"new" memory test that included targets and similar lures but not on an "old"/"new" memory test with targets and unrelated foils.

Kirwan et al. (2012) used a similar approach to investigate three patients with relatively selective hippocampal damage. Participants performed a memory test in which they viewed targets, similar lures, and unrelated foils. Participants were instructed to indicate whether items were "old" (i.e., exact repetition), "similar" (i.e., similar to—but not exactly the same as—a previously viewed image), or "new." They found that the patients were indistinguishable from controls in their ability to discriminate between previously viewed items and unrelated foil items. In contrast, the patients were impaired relative to controls at correctly responding "similar" to similar lures, supporting the findings from the patient in Holdstock et al. (2002). Duff et al. (2012) showed that patients with hippocampal lesions were impaired on a memory task when the stimuli were similar but performed as well as controls when the stimuli were dissimilar. Taken together, these results suggest that hippocampal damage impairs performance on tasks that require fine-grained discrimination between previously viewed stimuli, while sparing the ability to discriminate between previously viewed stimuli and unrelated foils. Altogether, these studies support the CLS models' predicted effects of hippocampal damage on recognition memory performance (Norman and O'Reilly, 2003; Norman, 2010).

Molecular Imaging and Electrophysiological Studies

Recent molecular imaging and electrophysiological studies have provided support for the role of the rodent hippocampus in pattern separation. For example, the immediate early gene, Arc, was sparsely activated in the DG following exposure to both novel and familiar environments (Chawla et al., 2005). Non-overlapping cells were activated in the DG when

rats were exposed to two different environments; however, the same cells were activated when rats were exposed twice to the same environment. Taken together, these results are consistent with pattern separation in the DG. Vazdarjanova and Guzowski (2004) investigated the overlap patterns of Arc activation in the CA3 and CA1 following exposure to contexts with varying degrees of similarity. Patterns of activity in CA1 linearly tracked the degree of change between environments. In contrast, the patterns of activity in CA3 were similar in response to very similar environments but were non-overlapping in response to two different environments. These results provide support for the prediction from computational models that suggest that CA3 contributes to pattern separation and pattern completion (e.g., Treves and Rolls, 1992, 1994; O'Reilly and McClelland, 1994; Hasselmo et al., 1996; Hasselmo and Wyble, 1997).

Leutgeb et al. (2004) used electrophysiology to investigate differences in patterns of activity in CA3 and CA1 as rats were exposed to environments that contained a varying degree of shared features. Similar to the immediate early gene study by Vazdarjanova and Guzowski (2004), the active cells in CA1 overlapped to the degree that the environments shared features. The active cells in CA3 were distinct across all of the environments, but the patterns of activity were very similar when an animal was re-exposed to the same environment. Taken together, these results support the role of CA3 in pattern separation. Leutgeb et al. (2007) recorded from the DG and CA3 as rats explored environments that were gradually morphed from a circular to a square enclosure (and vice versa). In the DG, the similarity between patterns of activity in response to two identical environments was significantly greater than the similarity between patterns of activity in response to two highly similar environments (i.e., $r(A, A) > r(A, A')$). In contrast, CA3 required a greater change in the shape of the environment before exhibiting different patterns of activity. Importantly, in EC, the patterns of activity were similar across the morphed environments, suggesting that the differences in patterns of activity in the DG were driven by local computations, rather than within DG afferents. These results suggest that both the DG and CA3 contribute to pattern separation,

but that the DG is more sensitive to small changes in the environment.

To investigate pattern separation in the DG and pattern completion in CA3, Neunuebel and Knierim (2014) incrementally rotated a circular maze. They found that small rotations were sufficient to elicit significantly different patterns of activity in the DG, whereas a larger rotation was required to significantly change patterns of activity in CA3. Therefore, the authors suggested that their results support the hypothesis that the DG performs pattern separation—i.e., small changes to the environment elicited reliably different patterns of activity—and that CA3 can perform pattern completion—i.e., stable patterns of activity across small changes to the environment). Taken together, there is evidence to suggest that the rodent hippocampus—in particular, the DG—plays a role in pattern separation. An important finding in these studies was the high degree of similarity between patterns of activity within the hippocampus (even the DG) in response to the same environments, suggesting that representations are stable in response to the same stimuli but that neurons in the hippocampus readily detect small changes to the environment.

Functional Magnetic Resonance Imaging Studies

Previous functional magnetic resonance imaging (fMRI) research has supported the hypothesized role of the human hippocampus in pattern separation (Kirwan and Stark, 2007; Bakker et al., 2008; Lacy et al., 2011; Motley and Kirwan, 2012). Before discussing these studies, a brief methodological background will be provided. Previous research has shown that several regions of the brain respond with a significant difference in blood-oxygen-level dependent (BOLD) activity in response to repeated presentations of an item compared to a first presentation of an item—an effect referred to as "repetition suppression." Technically, the way that it is defined here is agnostic to the direction of change, and the strict definition would be greater activity to first presentations than to repeated presentations; however, there is evidence that the direction of BOLD signal change might not always be the same as the di-
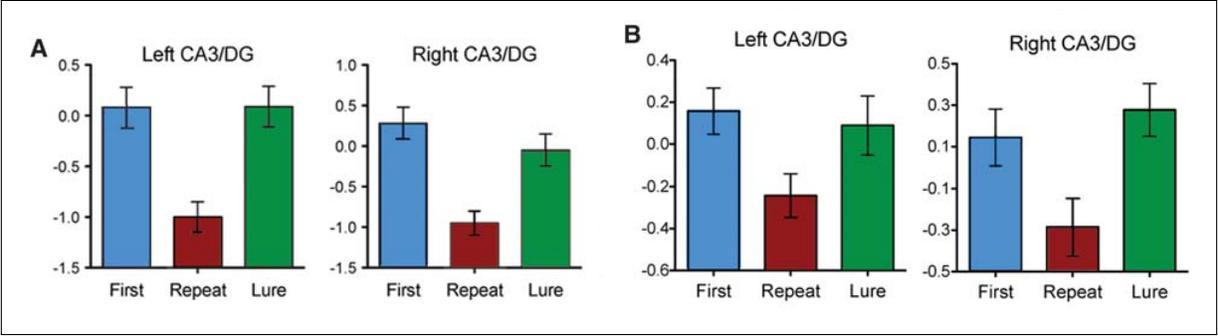
Figure 1.6: Pattern-separation-like activity in the human hippocampus (figure from Lacy et al., 2011). Repetition-sensitive voxels within the left and right DG/CA3 exhibited activity to related lures that was more similar to first presentations than it was to repeat presentations. Figure A shows results from Bakker et al. 2008 while Figure B shows replication in a novel dataset in Lacy et al. 2011.

rection of the local field potential change, which is an electrophysiological measure of neural activity (Hargreaves et al., 2012). Furthermore, there is no reason to assume that decreases in activity are less meaningful to the brain than increases in activity; hence, the preferred definition of repetition suppression is agnostic to the direction of change.

Bakker et al. (2008) scanned participants while they performed an incidental memory task. There were three trial conditions: 1) first presentations, 2) repeat presentations ("Repeat"), and 3) related lure presentations ("Lure"). A "First" versus "Repeat" contrast was used to select repetition-sensitive voxels within the MTL. Then the activity of "Lure" trials was compared to "First" and "Repeat" presentations. The criteria for pattern-separation-like activity were: 1) activity to "Lure" presentations that was significantly different than "Repeat" presentations, 2) activity to "Lure" presentations that was not significantly different than "First" presentations. The logic was that regions that respond with novel-like activity to related lures were treating these images as novel.

As predicted by rodent studies and computational modeling theories, Bakker et al. (2008) and Lacy et al. (2011) reported pattern-separation-like activity in the human DG/CA3 (see Figure 1.6). The results of Bakker et al. (2008) provided the first fMRI support of the role of the human hippocampus in pattern separation. An important limitation of the study by

Bakker et al. (2008) was that the authors did not manipulate the degree of similarity of the lure pairs. To address this limitation, Lacy et al. (2011) used a behavioral experiment to classify the lure pairs into high similarity and low similarity pairs. As discussed previously (see Figure 1.4), regions performing pattern separation would result in non-linear changes (i.e., greater than linear changes) to the output (see also "Model predictions" inset of Figure 1.7). Lacy et al. (2011) reported
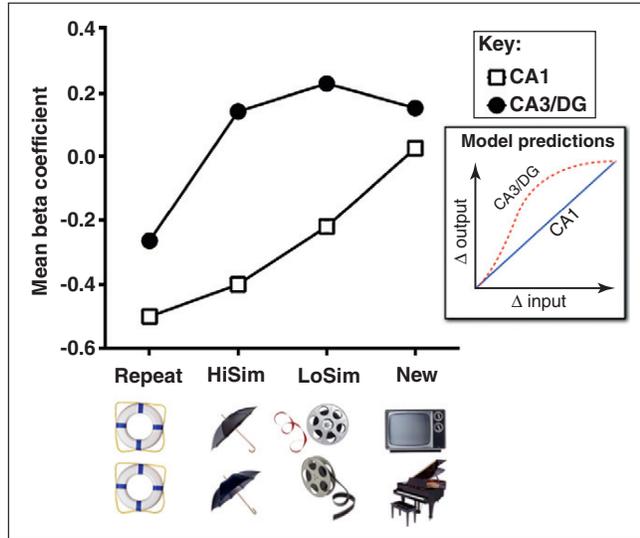
Figure 1.7: The DG/CA3 exhibited greater pattern-separation-like activity than region CA1 in response to highly similar lures (figure from Yassa and Stark, 2011).

that DG/CA3 exhibited greater pattern-separation-like activity than CA1 in response to high similarity lures, while DG/CA3 and CA1 responded similarly in response to low similarity lures (see Figure 1.7; Yassa and Stark, 2011 reanalyzed the data from Lacy et al., 2011). These findings support the hypothesis that the DG and CA3 are more biased toward pattern separation than CA1. Furthermore, these results are concordant with the hypothesis that the DG creates unique memory traces, even for similar stimuli (Marr, 1971; McNaughton and Morris, 1987; Treves and Rolls, 1992, 1994; O'Reilly and McClelland, 1994; Norman and O'Reilly, 2003; Yassa and Stark, 2011). These findings extend the results from the rodent hippocampus (Leutgeb et al., 2004, 2007; Vazdarjanova and Guzowski, 2004; Chawla et al., 2005; Neunuebel and Knierim, 2014) to the human hippocampus. Other authors have suggested that multivariate pattern analysis could provide more direct evidence for hippocampal pattern separation than activation analysis (Kumaran and Maguire, 2009; Bonnici et al., 2012), which I will discuss after providing a more thorough introduction to analytic techniques for fMRI.

1.3 An Introduction to Functional Magnetic Resonance Imaging

1.3.1 Activation Analysis

Early fMRI studies attempted to validate findings from neuropsychology—i.e., researchers used activation analysis in an effort to localize cognitive processes. In activation analysis, activation maps are compared between conditions (for an early example see: Petersen et al., 1988), thus revealing regions that show differences in activity for one condition compared to another. Activation analysis is typically a two-step process. First, univariate analysis is performed on every voxel—which is a box of fMRI data—in each participant independently. Second, the resultant statistical maps are warped to a common template to perform group-level analysis. This type of analysis makes two assumptions about cross-participant alignment: 1) the alignment to template space provides reasonable structural alignment, 2) the functional topography is similar. The first assumption is largely an empirical question (e.g., Yassa and Stark, 2009), and anatomical alignment techniques have become increasingly more accurate in recent years (e.g., Avants et al., 2008). The second assumption is largely theoretical and is more difficult to empirically validate. For example, it is possible that the hippocampus responds with greater activity to stimuli that will later be recollected than to those that will not. However, it is possible—perhaps likely—that the locus of activity will vary across participants. Most applications of activation analysis would not allow for this possibility. Region of interest (ROI) analysis—in which activity is averaged across the entire ROI—could detect activation regardless of location within the ROI as long as there were enough voxels with differential between-condition activity. However, as will be covered in the next section, ROI-based activation techniques could fail to detect between-condition differences in regions that represented information via distributed representations rather than overall differences in activity (e.g., Figure 1.8).

1.3.2 Multivariate Pattern Analysis

Activation analysis, such as the subtraction method (Petersen et al., 1988; Friston et al., 1996), is used by researchers that are attempting to find areas of the brain that are more active for one task or condition than another. It is certainly possible that brain regions operate by increasing their activation for certain tasks, however this is a relatively constrained view of brain function. If the brain is composed of regions of localized function, then it would make sense that localized regions would become active when their functions are occurring. As mentioned above, activation analysis was the dominant analytic framework in early fMRI studies. For example, Kanwisher and colleagues argued that there were face-selective regions and scene-selective regions of the brain—namely, the fusiform face frea (FFA; Kanwisher et al., 1997, 1999) and the parahippocampal place area (PPA; Epstein and Kanwisher, 1998; Epstein et al., 1999), respectively. They found that FFA responded maximally to faces and that PPA responding maximally to scenes.

As an alternative to the localization framework, it has been suggested that the brain represents information via distributed patterns of activity, similar to the concepts of psychological similarity and distributed representations discussed above. For example, extending Shepard's "second order isomorphism" hypothesis (Shepard and Chipman, 1970) from psychological space to the brain, the similarity between the patterns of brain activity in response to two stimuli should be similar to the psychological similarity between the two stimuli. For example, one should be able to find regions of the brain that represent—via distributed patterns of activity—two images of faces as more similar than an image of a face and an image of a scene, regardless of the overall activity in those regions (See Figure 1.8). In contrast, the localization framework attempts to find locations of the brain in which virtually all of the voxels are "on" (or show greater activity) to one stimulus category. Accordingly, the analysis of distributed patterns of activity can theoretically uncover more fine-grained information than activation analysis.
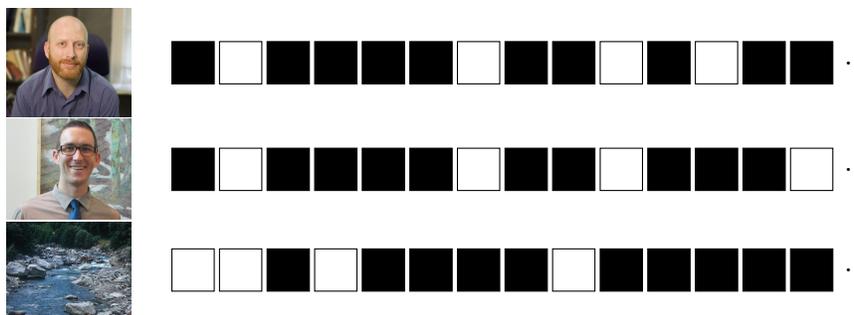
Figure 1.8: Representationally categorical patterns of activity. Each square represents a unit, and white indicates that a unit is "on." The patterns of activity across the units to the images of faces are very similar while the patterns of activity between either face image and the scene image are dissimilar. The patterns need not be non-overlapping between categories, e.g., the second unit is active to all three images. Activation-based analysis would fail in this example because the number of active units is similar between the two categories. Conversely, multivariate methods allow fine-grained patterns, such as the ones shown here, to be discovered.

Cutzu and Edelman (1998) used computational models to demonstrate that object and face representation could be achieved by coarsely-tuned overlapping patterns of activity, in line with the second order isomorphism hypothesis. Units in their models responded to several features, and the distributed pattern of activity across the units coded differences in the input patterns. The important theoretical distinction to the prevailing theory of the time (e.g., Kanwisher et al., 1997, 1999; Epstein and Kanwisher, 1998; Epstein et al., 1999) was that these models did not require distinct object or face processing units. Instead, object and face representation were achieved based on the similarity of activity across the units. Edelman's computational models were similar to some of the previously mentioned models (e.g., Rumelhart and Todd, 1993). Edelman et al. (1998) validated their computational models by studying distributed patterns of activity in the human brain. They performed an fMRI experiment in which participants viewed stimuli from several categories. They selected object-responsive voxels—greater activation to objects than baseline—in the ventral visual stream. Then, they calculated the Euclidean distance between patterns of activity across the selected voxels to several categories of objects, which provided an index of similarity between the patterns of activity in response to different stimuli, such that stimuli that elicited similar

patterns of activity—i.e., smaller Euclidean distances between patterns of activity—were inferred to have been represented similarly. The results from their analysis showed that stimuli that were deemed to be behaviorally similar were represented similarly in the brain. Their paper was important for three reasons. First, it was the first fMRI paper to investigate psychological and representational similarity—i.e., it was the first paper to test the second order isomorphism hypothesis in the human brain. Second, it was the first paper to use fMRI data to compare distributed representations in computational models and brain activity, thus providing a link between computational theory and neuroimaging analysis. Third, it was the first neuroimaging multivariate pattern analysis paper, which was a fundamentally different approach to thinking about how the brain represents information.

Haxby et al. (2001) is often cited as the first multivariate pattern analysis (MVPA) study (cf. Haxby, 2012). Haxby et al. (2001) analyzed patterns of activity in the ventral temporal lobe in response to 8 categories of visual stimuli. Using a correlational classifier, they showed successful between-category classification accuracy. They found that even when removing the most activation-based responsive voxels (i.e., the PPA and FFA), there still was significant category representation of all 8 categories. They also found that when they investigated category representation within regions that responded maximally to one category (e.g., the FFA for faces), there was significant between-category classification for the other 7 categories. This suggests that while the FFA and PPA may respond maximally to one category, these regions also carry information about other stimulus categories. The definition of a category-selective region is that it responds maximally to one stimulus category and does not differentiate between other stimulus categories; hence, the FFA and PPA are perhaps preferentially tuned to processing faces and scenes, respectively, but they are not, in fact, category selective. The results of Haxby et al. (2001) are theoretically important because they suggest that brain regions can represent stimuli differently without changing overall levels of activation, thus obviating the need for object and face processing modules. Instead, their results suggest that the brain represents categories by the similarity of the resultant

patterns of activity; hence, theoretically unlimited categories could be represented without the need for unlimited modules.

The second order isomorphism framework can be extended to between-participant differences in representations. It is possible that all participants have a pattern of activity that is more face-like or more scene-like, but the actual activity pattern is unique to each participant. Rather than assuming functional alignment across participants as does activation-based analysis, the majority of MVPA studies analyze activity in individual participants. Measures such as classification accuracy can then be compared across participants, thus only assuming that the representation of stimuli is similar across participants in the studied regions.

Kriegeskorte et al. (2008a) used similar techniques to Edelman to create a subcategory of MVPA called Representational Similarity Analysis (RSA). While Edelman et al. (1998) used Euclidean distance as a measure of distance between patterns of activity, Kriegeskorte has used Pearson's correlation coefficient for a measure of similarity or correlation distance (1-Pearson's correlation coefficient) as a measure of distance. Pairwise correlations are entered into a correlation matrix, in which each entry in the matrix represents the similarity between patterns of activity in response to two different stimuli. Kriegeskorte has argued that the correlation matrix allows abstraction away from the data itself into a similarity space, thus RSA provides a framework for easily comparing representations across individuals, across species, and across computational models (Kriegeskorte et al., 2008a,b; Kriegeskorte, 2009). For example, the between-species comparison of representational similarity matrices suggested that the representation of objects in the inferior temporal cortex was very similar in monkeys and humans (Kriegeskorte et al., 2008b). Interestingly, the human data were acquired using high-resolution fMRI while the monkey data were acquired using intracranial electrophysiology. Furthermore, group-level power can be obtained by combining RSA matrices, without assuming functional alignment across subjects (Kriegeskorte et al., 2008a).

Another highlight of RSA is that it does not assume a pre-defined representational struc-

ture (Kriegeskorte et al., 2008a; Kriegeskorte, 2009). For example, many studies employing classification analysis use stimuli that come from distinct categories (e.g., faces and objects). Then researchers typically investigate whether patterns of activity within brain regions can successfully classify between the stimulus categories. This provides valuable information as to whether the investigated brain regions exhibit category representation; however, it does not provide information about within-category representation. Within-category similarity provides useful information—i.e., it could be the case that the within-category similarity differs across brain regions without affecting the overall between-category similarity (Kriegeskorte et al., 2008a; Kriegeskorte, 2009). For example, a brain region could exhibit significant classification accuracy for faces versus objects with relatively similar patterns of activity to pictures of faces and relatively dissimilar patterns of activity between pictures of objects. Thus, investigation of within-category similarity provides additional information, and RSA naturally allows such representations to be uncovered (Kriegeskorte et al., 2008a; Kriegeskorte, 2009).

Kriegeskorte et al. (2008a) described a novel analytic technique, representational connectivity, which allows comparison of representations across ROIs. Representational connectivity is particularly interesting for areas that are anatomically connected as it can provide insight into the computations that may take place in different regions of the brain by examining transformations of representations. Representational connectivity provides a static measure of the similarity of representations across different regions. In contrast, Coutanche and Thompson-Schill (2013) described a novel analytic technique, informational connectivity, which investigates time courses of representation across different regions. Informational connectivity provides the link between functional connectivity and static representational connectivity (Kriegeskorte et al., 2008a) by considering how representations correlate with each other over time across different brain regions.

Classification analysis has also been used to investigate the similarity structure of representa-

tions (e.g., Walther et al., 2009, 2012). Confusion matrices can be generated when performing classification analysis. Confusion matrices provide information about the between-stimulus similarity (based on classification errors). In the case of perfect classification, the confusion matrix would be the identity matrix (1's in the diagonal and 0's everywhere else). In the case of imperfect classification, off-diagonal values represent classification errors. Theoretically, stimuli that are represented similarly should result in a higher number of classification errors between those stimuli. For example, in a study that contains categorical images, a higher number of within-category classification errors would be predicted (because similar items are more confusable). Behavioral experiments have been used for decades to investigate the psychological similarity of stimuli. More recently, studies have used confusion matrices derived from behavior and patterns of activity in the brain to find regions that show significant similarity between behavioral and neural representation and thus could theoretically be involved in behavioral decisions (Walther et al., 2009, 2012).

Classifier-based confusion matrices are theoretically identical to the correlation matrices of RSA; however, one benefit of classification techniques, such as linear support vector machines (SVM), is that they typically contain regularization techniques (Davis and Poldrack, 2013), whereas RSA does not. Accordingly, SVM analysis has been shown to perform well under conditions with many voxels, including many uninformative voxels (Etzel et al., 2009). In effect, SVM can account for uninformative voxels by assigning them very low weights. Taken together, investigation of SVM-based confusion matrices can partially correct for between-region differences in noise and the number of uninformative voxels across regions, and hence might be more robust than RSA (cf. Davis and Poldrack, 2013). However, each stimulus must be presented several times to obtain reliable within-stimulus classification accuracy. In any case, RSA and classification analysis provide a principled framework for the investigation of representation.

A theoretical problem for MVPA is whether decoded brain patterns are actually used by

the brain or the animal in a behaviorally meaningful manner. Pais-Vieira et al. (2013) used related techniques to analyze electrophysiological data in rodents. They discovered that decoded patterns of activity are behaviorally meaningful. Additionally, they elicited task-relevant behavior upon stimulating the brain with the patterns of activity that were decoded during the performance of a cognitive task, suggesting that decoded patterns of activity may indeed be useful to the brain and behavior. Recent optogenetic experiments in mice have revealed similar results (Cowansage et al., 2014; Tanaka et al., 2014). Taken together, animal studies can provide more causal evidence for the involvement of distributed patterns of activity in behavioral performance. Brain-behavior correlations can also be useful for examining the involvement of regions in behavioral performance. For example, as mentioned above, fMRI studies of human participants can investigate whether there is a relationship between representations in various brain regions and behavioral performance; however, the results of correlation analysis will not speak to whether regions play a causal or necessary role. Accordingly, converging evidence from animal and human studies will be necessary to clarify whether decoded representations are related to behavior.

Summary of Multivariate Pattern Analysis

Taken together, MVPA provides a framework for examining representations across data acquisition techniques. MVPA provides an excellent platform for constraining computational models and theories of the brain because it provides a coherent framework for comparing model and brain representations. MVPA can be used to study representations in the brain as a whole (whole brain MVPA), representations within regions of interest, and transformations of representations across anatomically connected regions. Because MVPA does not assume widespread activity changes, it has the potential to guide new findings of the brain that are closely related to findings from computational models of brain function, e.g., distributed representations. In contrast to activation analysis that assumes localized activity changes

within specific brain regions, MVPA assumes that the brain represents the world by the similarity (and dissimilarity) of patterns of activity.

1.3.3   Pattern Separation in the Human Hippocampus, Revisited

As mentioned above, previous authors have suggested that MVPA could be used to more directly test the involvement of the human hippocampus in pattern separation (Kumaran and Maguire, 2009; Bonnici et al., 2012), and recent studies have used MVPA to investigate patterns of activity within the MTL. Diana et al. (2008) performed a high-resolution fMRI experiment aimed at investigating category representation. They showed significant classification accuracy between images from several categories using patterns of activity in PHC. Conversely, they found that classification accuracy in the hippocampus was not different than chance for any of the investigated categories.

In a similar experiment, Liang et al. (2013a) observed significant classification accuracy between images from several categories in PHC and PRC. They also found significant classification accuracy for scenes compared to other categories within the posterior hippocampus but not in the anterior hippocampus. They suggested that it was possible that Diana et al. (2008) failed to find a difference in classification accuracy between scenes and other categories within the hippocampus because they did not separate the hippocampus into posterior and anterior subregions. Given that Diana et al. (2008) used classification techniques that included regularization parameters, it would seem that if posterior hippocampus was carrying information capable of distinguishing between scenes and other categories, then the voxels in posterior hippocampus would have been weighted strongly while those in the anterior hippocampus would have been weighted very weakly (due to low information in that subregion). In effect, it would seem that Diana et al. (2008) would have found significant classification in the hippocampus if posterior hippocampus was carrying information, therefore the fact

that they did not suggests that there might be minor differences between the two findings. In any case, Liang et al. (2013a) showed that posterior and anterior hippocampus showed lower classification accuracy than PHC and PRC.

LaRocque et al. (2013) investigated category representation within PHC, PRC, and the hippocampus using RSA. They provided evidence that the MTL cortical regions—but not the hippocampus—were representationally categorical. They also observed a significant positive relationship between traditional recognition memory and within-category similarity in PHC and PRC, and a negative relationship between traditional recognition memory and within-category similarity in the hippocampus. These results suggest that higher within-category similarity in PHC and PRC is beneficial to performance, while lower within-category similarity in the hippocampus is beneficial for performance. The latter result supports the notion that hippocampal pattern separation plays a role in memory performance. Taken together, their results provide evidence for behavioral relevance of representations in PHC, PRC, and the hippocampus.

We recently performed an experiment to investigate representations within the medial temporal lobe (Huffman and Stark, 2014; see **Chapter 2**). In our report, we coined the term "representationally agnostic" to refer to brain regions that are not representationally categorical, with the added criterion that they show task-dependent modulation. We hypothesized that the hippocampus would be more representationally agnostic than MTL cortical regions due to its pattern-separated representations, and we hypothesized that MTL cortical regions would be representationally categorical due to their hypothesized ability to compute summed similarity (Norman and O'Reilly, 2003; Norman, 2010). Pattern separation tends to assign non-overlapping representations, even in response to similar stimuli. Such a representational scheme would result in low within-category similarity, which would result in low between-category classification accuracy (see Figure 1.9; cf. Chadwick et al., 2012). I will refrain from discussing the results of our experiments until the next chapter, but I will mention
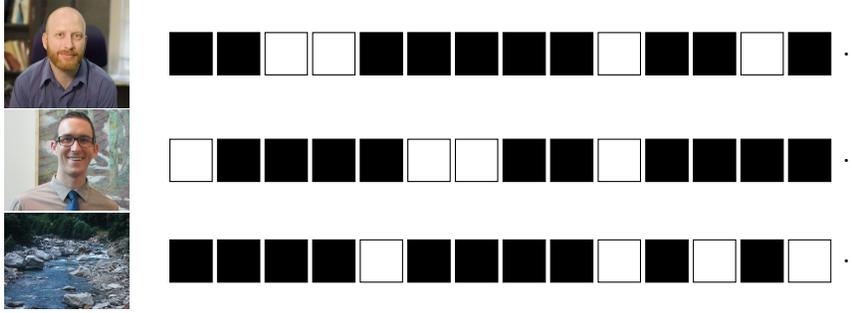
Figure 1.9: Representationally agnostic patterns of activity. White indicates that a unit is "on." Pattern separation would result in representationally agnostic patterns of activity because it would create unique patterns of activity even for similar stimuli—i.e., the two images of faces are represented as dissimilarly as the images of faces and the image of a scene.

that Diana et al. (2008), Liang et al. (2013a), and LaRocque et al. (2013) have provided indirect support for the role of the human hippocampus in pattern separation by providing evidence that the hippocampus is more representationally agnostic than MTL cortical regions. Therefore, these studies have extended upon previous results from activation analysis of fMRI data in the human hippocampus (Kirwan and Stark, 2007; Bakker et al., 2008; Lacy et al., 2011; Motley and Kirwan, 2012).

More direct evidence of pattern separation in the human hippocampus has been provided by several recent papers that have used MVPA to investigate patterns of activity in the hippocampus. Specifically, there have been reports of significant classification accuracy in the hippocampus between distinct locations in a virtual environment (Hassabis et al., 2009; but see: Op de Beeck et al., 2013), between memories of highly overlapping video clips (Chadwick et al., 2010, 2011), and between highly similar images of scenes (Bonnici et al., 2012). Additionally, these studies reported significantly greater classification accuracy in the hippocampus than in MTL cortical regions. Taken together, the studies reviewed in this section provide support for the predictions of the CLS model (Norman and O'Reilly, 2003; Norman, 2010); specifically, MTL cortical regions appear to exhibit similar patterns of activity in response to similar stimuli while the hippocampus is biased toward pattern

separation, showing more distinct patterns of activity in response to similar stimuli than the MTL cortical regions.

1.4   Overall Conclusions

Theoretical accounts have suggested for many years that memory is not a single entity; however, neurobiological studies were required to provide more direct evidence that memory systems could be independent. Studies of patient H.M. provided evidence for a single dissociation, while subsequent animal research provided evidence for a double dissociation. Altogether, studies of patients with MTL damage and animal investigations suggested a dissociation between MTL-dependent declarative memory and MTL-independent procedural memory. Dual-process theories suggested that recollection and familiarity independently contribute to recognition memory performance. Research employing the lesion method, electrophysiology, and neuroimaging suggested that there these processes differentially rely on the hippocampus and PRC; however, other theories suggested that regions of the MTL are more broadly involved in declarative memory. In contrast to differences of opinion regarding the localization of recollection and familiarity, proponents of both views have hypothesized that subregions of the MTL differ in terms of the information that they represent. Specifically, PHC is hypothesized to be important for memory for scenes and contexts while PRC is important for memory for objects. Moreover, recent theories suggest that retrosplenial cortex also plays a critical role in declarative memory. We used these frameworks as our overarching hypothesis for our fMRI experiments.

In this dissertation, we describe the results of four fMRI studies, in which we used multivariate pattern analysis to investigate the representation of information in the MTL and retrosplenial cortex. Our first aim was to investigate whether parahippocampal cortex and retrosplenial cortex are preferentially involved in the processing of scene information (see

Chapter 2). Additionally, the CLS model predicts that the MTL cortex would be able to compute the summed similarity of the stimuli while the hippocampus would not due to differences in the bias toward pattern separation in these regions (Norman and O'Reilly, 2003; Norman, 2010). We also tested the summed similarity hypothesis using fMRI and multivariate pattern analysis (see **Chapter 2**). Our second aim was to investigate the representation of individual items and contexts in the medial temporal lobes and retrosplenial cortex (see **Chapter 3**).

Recent evidence has suggested that healthy older adults show a disproportionate impairment on tasks that tax associative memory compared to item memory (Craik and McDowd, 1987; Spencer and Raz, 1995; Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2004; Old and Naveh-Benjamin, 2008a,b; Danckert and Craik, 2013). An alternative, but not mutually exclusive, hypothesis is that healthy older adults encode fewer features during learning. We tested the latter hypothesis using a combined behavioral and model-based approach. Specifically, our third aim was to investigate memory performance on a forced-choice variant of the Mnemonic Similarity Task in healthy younger adults and the changes that occur as a result of healthy aging. Additionally, we used global matching models to provide a potential mechanistic account of the empirical results (see **Chapter 4**). In the final chapter of this dissertation, I discuss the results of our functional, behavioral, and model-based experiments and provide suggestions for future research (see **Chapter 5**).

Chapter 2

Scene representation in parahippocampal cortex and retrosplenial cortex

The following chapter was previously published in *Hippocampus* (Huffman and Stark, 2014). Contemporary theories of the medial temporal lobe (MTL) suggest that there are functional differences between the MTL cortex and the hippocampus. High-resolution functional magnetic resonance imaging and multivariate pattern analysis were utilized to study whether MTL subregions could classify categories of images, with the hypothesis that the hippocampus would be less representationally categorical than the MTL cortex. Results revealed significant classification accuracy for faces versus objects and faces versus scenes in MTL cortical regions—parahippocampal cortex and perirhinal cortex—with little evidence for category discrimination in the hippocampus. MTL cortical regions showed significantly greater classification accuracy than the hippocampus. The hippocampus showed significant classification accuracy for images compared to a non-mnemonic baseline task, suggesting that it responded to the images. Classification accuracy in a region of interest encompassing retrosplenial cortex and the posterior cingulate cortex posterior to retrosplenial cortex (RSC/PCC), showed a similar pattern of results to parahippocampal cortex, supporting the hypothesis that these regions are functionally related. The results suggest that parahip-

pocampal cortex, perirhinal cortex, and RSC/PCC are representationally categorical and the hippocampus is more representationally agnostic, which is concordant with the hypothesis of the role of the hippocampus in pattern separation.

2.1 Introduction

The Complementary Learning Systems (CLS) neural-network model posits that fundamental computational trade-offs have led the brain to develop multiple memory systems. In particular, it advances the notion that one system—or set of systems—is specialized for extracting statistical regularities of the world through gradual, interleaved learning and another is specialized for rapid, arbitrary associative learning via pattern separation (McClelland et al., 1995; O'Reilly and Rudy, 2000; O'Reilly and Norman, 2002; Norman and O'Reilly, 2003; Norman, 2010). Pattern separation transforms overlapping input patterns into more dissimilar patterns, and hence could theoretically allow rapid learning of novel information without high levels of interference to existing, potentially similar memories. Using the CLS framework, Norman (2010) posited that patterns of activity in the MTL cortex (MTLC) allow computation of summed similarity (i.e., similar stimuli elicit similar patterns of activity) while the hippocampus cannot compute summed similarity due to its pattern-separated representations (i.e., similar stimuli elicit unique patterns of activity). Neither form of learning or representation alone would provide an adaptive memory system, leading to the need for multiple memory systems operating under different computational principles.

Rodent studies provided neural (Leutgeb et al., 2005, 2007; Leutgeb and Leutgeb, 2007) and behavioral (Gilbert et al., 1998; Potvin et al., 2009; Talpos et al., 2010; Morris et al., 2012) evidence supporting the hippocampus' role in pattern separation. Supporting the CLS model's predictions of behavioral changes caused by hippocampal lesions (Norman and O'Reilly, 2003; Norman, 2010), investigations of patients with selective hippocampal damage

revealed deficits in behavioral pattern separation tasks, with relative sparing of traditional recognition memory (Holdstock et al., 2002, 2005; Mayes et al., 2002; Duff et al., 2012; Kirwan et al., 2012; but see Reed and Squire, 1997; Stark and Squire, 2003; Bayley et al., 2008). Additionally, recent functional magnetic resonance imaging (fMRI) studies have shown activity signals consistent with pattern separation in the human hippocampus (e.g., Bakker et al., 2008; Lacy et al., 2011; Motley and Kirwan, 2012; for across-species review of pattern separation see Yassa and Stark, 2011). These studies leveraged the repetition-suppression effect, in which different levels of brain activity are observed to a first presentation of a stimulus compared to a repeated presentation of a stimulus, to assess whether activity in different regions was consistent with pattern separation. The hippocampus exhibited first-presentation-like activity levels in response to related lures, suggesting that the hippocampus treated related lures as novel images, thus supporting the hypothesized role of the human hippocampus in pattern separation.

Multivariate pattern analysis (MVPA; for reviews see Haynes and Rees, 2006; Norman et al., 2006) investigates spatial patterns of activity across voxels. Depending on experimental design, MVPA can be used to analyze patterns of activity at the individual trial level (e.g., slow event-related design) or at the level of several within-category stimuli (e.g., blocked design). Patterns of activity can be subjected to machine learning classification analysis to uncover whether a brain region—or brain regions—exhibits different patterns of activity in response to different categories of stimuli. Pattern separation is defined computationally as the formation of output patterns of activity that are less similar than input patterns of activity. Because MVPA investigates patterns of activity, it can be used to more directly test the hypothesis of hippocampal pattern separation than univariate techniques (cf. Chadwick et al., 2012). For example, recent studies have shown greater classification accuracy between similar stimuli in the hippocampus than the MTLC, thus suggesting that the hippocampus forms more distinct representations of similar stimuli (Hassabis et al., 2009; Chadwick et al., 2010, 2011; Bonnici et al., 2012). MVPA has also been used to investigate category

representation in the MTL (Diana et al., 2008; Liang et al., 2013a; LaRocque et al., 2013).

We refer to brain regions that show significant classification accuracy between categories of images as being representationally categorical, and we coin the term representationally agnostic to refer to brain regions that are not representationally categorical. Providing evidence that the MTLC is representationally categorical would support its hypothesized ability to compute summed similarity because it would suggest that similar items (items within a category) are represented similarly (with distinct representations across categories), while providing evidence that the hippocampus is far more representationally agnostic (i.e., category classification accuracy that is consistently at or near chance) would support its hypothesized role in pattern separation. Accordingly, MVPA can be used to test the hypothesis that the MTLC can compute summed similarity while the hippocampus cannot.

Previous studies suggested that the MTLC is more representationally categorical than the hippocampus (Diana et al., 2008; Liang et al., 2013a; LaRocque et al., 2013). Two of these studies used blocked designs (Diana et al., 2008; Liang et al., 2013a), thus precluding the study of representations of individual stimuli. LaRocque et al. (2013) used a rapid event-related design, presenting the same stimuli in each run, thus allowing use of the general linear model to extract stimulus-specific estimates of activity. Here we report the results of two experiments in which we used a slow event-related design to allow for the extraction of patterns of activity at the level of individual trials. Using MVPA, we aimed to replicate previous results using individual trial patterns of activity and to rule out the possibility that the hippocampus is more representationally agnostic than the MTLC due to: 1) lack of responsiveness to the stimuli, 2) differing numbers of voxels across the MTLC and the hippocampus, or 3) noisier data.

In addition to testing this central hypothesis, we examined category representation within a region of interest encompassing retrosplenial cortex (RSC; traditionally defined as Brodmann areas 29 and 30; Vann et al., 2009) and the posterior cingulate cortex (PCC) posterior to

RSC. Recent accounts have suggested that RSC should be considered as part of the network of regions involved in memory (Vann et al., 2009; Ranganath and Ritchey, 2012). Specifically, Ranganath and Ritchey (2012) posited that there are two cortical systems involved in memory—one in the anterior temporal lobe that includes PRC, and another in the posterior cortex that includes PHC, RSC, and PCC. Several recent accounts have suggested that PHC, RSC, and PCC are anatomically situated to be involved in representing contextual information, including scenes or spatial information (Buffalo et al., 2006; Vann et al., 2009; Wixted and Squire, 2011; Ranganath and Ritchey, 2012). We predicted that RSC/PCC would show a similar pattern of classification results to those in PHC. Finally, we performed a variant of informational connectivity (Coutanche and Thompson-Schill, 2013) in which we investigated the correlation between the trial-by-trial multivariate pattern discriminability in PHC and RSC/PCC. Given the predicted role of these regions in scene and spatial representation, we hypothesized that the multivariate pattern discriminability in these regions would be correlated during an experiment that contained images of scenes.

2.2 Methods and Materials

2.2.1 Participants

A total of 29 participants (12 males, 17 females) underwent MRI scanning. Three were excluded because half or fewer of the total functional runs were obtained, one was excluded for sleeping during functional scans, and one was excluded due to excessive motion; therefore, 24 participants were used for analysis, with 12 in each of two experiments. Participants consented to the procedures in accordance with the Institutional Review Board of the University of California, Irvine and received \$25 for their first hour of participation and \$5 for each additional 20 minutes.

2.2.2 Stimuli

In Experiment 1, stimuli consisted of face and object images (12 each per run, totaling 192 each over 16 runs). In Experiment 2, stimuli consisted of face and scene images (8 each per run, totaling 128 each over 16 runs). Face images in both experiments included the hair and neck of the depicted individual. Half of the stimuli in each run were novel images, one quarter were repeated images, and the remaining quarter were related lure images. The novel versus repeated status of images was irrelevant for the present experiment. Participants viewed images and were asked to indicate, using an MRI-compatible button box, whether they found the picture to be pleasant or unpleasant. Each picture was displayed for 4 seconds, followed by an 8 second interstimulus interval (ISI). To discourage mind wandering during the ISI, subjects performed an engaging perceptual but non-mnemonic arrows task (Stark and Squire, 2001), in which they indicated the direction (left or right) in which arrows were pointing on the screen (6 trials per ISI; the ISI was modeled after Kuhl et al., 2011, 2012).

In Experiment 2, subjects also performed an engaging, but non-mnemonic perceptual baseline (PB) task (Law et al., 2005; Mattfeld and Stark, 2010, 2011; Hargreaves et al., 2012; 8 per run, totaling 128 for 16 runs). In the current version, subjects made two subsequent judgments (totaling 4 seconds) about which of two boxes was brighter (on a random static background), followed by the arrows task during the 8 second ISI. To keep the perceptual baseline task challenging, the difference in brightness of the boxes was continuously titrated to keep subjects between 50 and 70 percent accurate.

2.2.3 Image Acquisition

Data were collected using a 3.0-T Philips scanner using a sensitivity encoding (SENSE) head coil at the Research Imaging Center at University of California, Irvine. A whole-brain 1.0 mm isotropic 3D magnetization-prepared rapid gradient echo (MP-RAGE) structural scan was

47

collected. Functional data were acquired using a T2*-weighted echo planar imaging sequence (TE=26 ms, TR=2400 ms, 32 slices, 1.5 mm isotropic voxels). To allow for magnetic field stabilization, the first 5 TRs of each run were immediately discarded. One hundred and twenty volumes were collected in each functional run. The full experiment consisted of 16 runs; however, for technical reasons data were obtained (or processed) for fewer than 16 runs for 4 subjects (mean number of runs=15.8, minimum=14).

2.2.4   Preprocessing

Functional data were aligned to the subjects MP-RAGE using AFNIs align_epi_anat.py script (Saad et al., 2009). Data were quadratically detrended, high-pass filtered ($f > 0.01$ Hz), and minimally smoothed by a 2 mm FWHM Gaussian kernel, using AFNI (Cox, 1996). Data were then z-scored and three TRs (beginning at 2.4, 4.8, and 7.2 seconds after stimulus onset) were averaged to form individual trial activity patterns, using PyMVPA (Hanke et al., 2009).

2.2.5   Regions of Interest

Anatomical regions of interest (ROI) for hippocampus, parahippocampal cortex (PHC), and perirhinal cortex (PRC) were manually segmented on a model template (see Figure 2.1A). PRC was labeled according to landmarks described by Insausti et al. (1998), and PHC was defined as the portion of the parahippocampal gyrus caudal to the perirhinal cortex and rostral to the splenium of the corpus callosum, as in our previous research (Stark and Okado, 2003; Kirwan and Stark, 2004; Law et al., 2005; Okado and Stark, 2005). Each participant's MP-RAGE was aligned to the model template, using Advanced Normalization Tools (ANTs; Avants et al., 2008), allowing the model template ROIs to be warped to each subjects original space using the inverse warp vectors. We utilized Freesurfer's isthmus cingulate
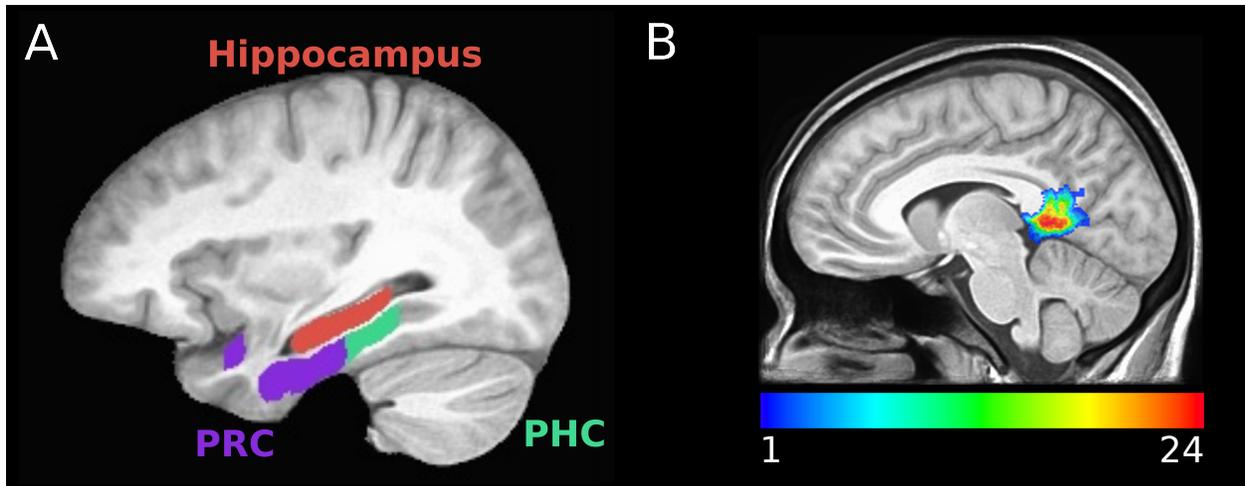
Figure 2.1: ROI locations. (A) Hippocampus (red), parahippocampal cortex (green), and perirhinal cortex (purple) within model space. (B) For visualization, RSC/PCC participant masks were warped to template space to create an overlap map of the ROI across participants (24 participants total).

label (Desikan et al., 2006)—which encompasses RSC (traditionally defined as Brodmann's areas 29 and 30; Vann et al., 2009) and PCC posterior to RSC—for our RSC/PCC ROI. The data were collected using a partial brain coverage acquisition box, resulting in slightly limited coverage of the superior aspect of RSC/PCC (see Figure 2.1B). Each participant's ROIs were resampled to functional resolution and masked to contain only completely sampled voxels (mean number of voxels: hippocampus=1803, PHC=1270, PRC=1888, RSC/PCC = 655).

2.2.6 Multivariate Pattern Analysis

Classification analysis was performed using PyMVPA (Hanke et al., 2009) and custom-written Python code (`python.org`), using the NeuroDebian package repository for Linux (Hanke and Halchenko, 2012). Linear support vector machine (SVM) analysis was carried out in each subjects original space, using individual trial patterns of activity within each ROI (see Figure 2.1). The classification accuracy across each N-fold-leave-one-run-out cross validation was averaged to generate an average classification accuracy for each participant for each ROI. Two-tailed one sample t-tests of group data were used to assess significance of classification

accuracy (compared to chance) and two-tailed paired t-tests were used to assess classification accuracy differences between ROIs. Resultant p-values were Bonferroni corrected for six comparisons (three ROIs and three comparisons between ROIs). Welch's t-tests, which can account for unequal variance between samples, were used to assess significance of differences between experiments. R (`cran.r-project.org`) was used to perform t-tests and to plot figures.

To test the hypothesis that RSC/PCC was representationally similar to PHC, we performed a variant of informational connectivity (Coutanche and Thompson-Schill, 2013). The original instantiation of informational connectivity utilized a correlational classifier to calculate the within minus between category similarity to create a vector of multivariate pattern discriminability on a TR-by-TR basis. Then the Spearman rank correlation between multivariate pattern discriminability in different regions (in their case a seed region and a spherical searchlight region) was calculated. Here we utilized SVM decision values (referred to as estimates in PyMVPA) on a trial-by-trial basis to test whether the multivariate pattern discriminability was correlated between PHC and RSC/PCC. To maximize similarity to the previously described method (Coutanche and Thompson-Schill, 2013), the value of multivariate pattern discriminability was set to be greater than zero when the classifier's prediction was correct and less than zero the classification prediction was incorrect. Spearman's rank correlation was calculated between the multivariate pattern discriminability vectors in PHC and RSC/PCC. Resultant values were Fisher's r to z transformed (z[r]).

2.3 Results

Linear SVM classification analysis was performed using individual trial patterns of activity from all of the voxels within each ROI. The stimulus set contained novel as well as non-novel images within each run (see Methods). The same pattern of results was obtained when using
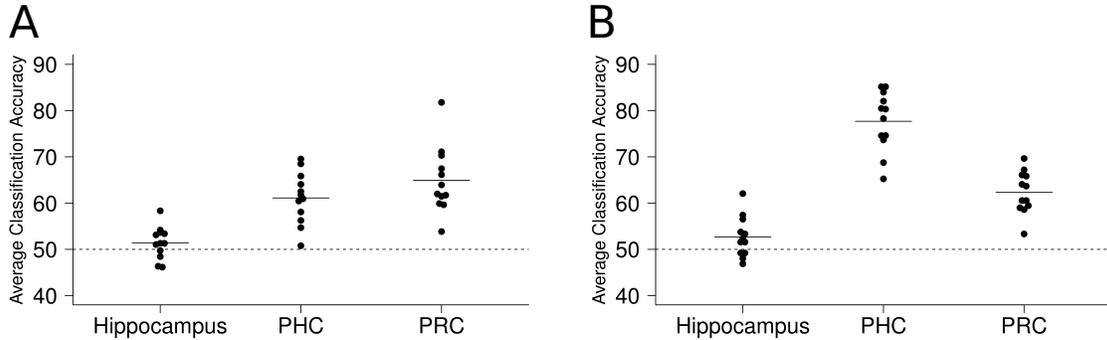
Figure 2.2: Linear SVM classification results. Each dot represents the average classification accuracy for one participant and each line represents the group ROI mean. (A) Experiment 1—faces versus objects. (B) Experiment 2—faces versus scenes. Classification accuracy was significantly greater than chance in MTLC but not in hippocampus, and was significantly greater in MTLC than in hippocampus. Classification accuracy for faces versus scenes was greater in PHC than PRC, and PHC classification accuracy was greater for faces versus scenes than faces versus objects.

the full stimulus set as when using the novel-only stimulus set, and paired t-tests revealed no significant differences between the two stimulus sets (all MTLC and hippocampal $t_{11} < 1.4$, uncorrected $p > 0.19$); therefore, we limit our discussion to the results obtained from the full stimulus set because it contained more stimuli.

In Experiment 1, one sample t-tests revealed significant classification accuracy for faces versus objects in PHC and PRC (mean classification accuracy: PHC $= 61.1\%$, PRC $= 64.9\%$; both $t_{11} > 6.9$, corrected $p < 0.001$) but not in the hippocampus (mean classification accuracy $= 51.4\%$; $t_{11} = 1.41$, uncorrected $p = 0.19$, see Figure 2.2A). Paired t-tests revealed significantly greater classification accuracy for faces versus objects in PHC and PRC than in the hippocampus (both $t_{11} > 4.9$, corrected $p < 0.005$), with no significant difference between PHC and PRC ($t_{11} = 1.86$ uncorrected $p = 0.09$, corrected $p = 0.53$).

In Experiment 2, one sample t-tests revealed significant classification accuracy for faces versus scenes in PHC and PRC (mean classification accuracy: PHC $= 77.7\%$, PRC $= 62.3\%$; both $t_{11} > 9$, corrected $p < 0.001$) with little evidence for above-chance category discrimination in the hippocampus (mean classification accuracy $= 52.7\%$; $t_{11} = 2.11$, uncorrected $p = 0.06$,

Figure 2.3: Linear SVM classification results of images versus the perceptual baseline task (Experiment 2). The hippocampus significantly classified faces versus perceptual baseline and scenes versus perceptual baseline.

corrected $p = 0.35$, see Figure 2.2B). Paired t-tests revealed significantly greater classification accuracy for faces versus scenes in PHC and PRC than in the hippocampus (both $t_{11} > 7$, corrected $p < 0.005$) and in PHC than PRC ($t_{11} = 5.8$, corrected $p < 0.001$). In contrast to classification results between the image categories, one-sample t-tests revealed significant classification accuracy for images versus the perceptual baseline task in the hippocampus (mean classification accuracy: faces vs PB = 53.9%, scenes vs PB = 57.1%; both $t_{11} > 3$, $p < 0.05$, corrected for 6 comparisons for consistency with the previous analyses; see Figure 2.3).

Bonferroni corrected (for 3 comparisons), two-tailed Welchs two sample t-tests were used to compare the classification accuracy results obtained from Experiment 1 and 2. This revealed significantly greater classification accuracy in PHC for faces versus scenes than faces versus objects ($t = 6.75$, corrected $p < 0.001$). Classification accuracy within PRC and within the hippocampus was not significantly different between Experiment 1 and 2 (both $t < 1.1$, uncorrected $p > 0.3$).

To rule out the possibility that the previous results were driven by differing numbers of voxels across ROIs (or due to a differing number of noisy voxels), voxel-wise signal to noise

(SNR; mean intensity/standard deviation) distributions were inspected and voxel selection was performed. The median SNR was similar across ROIs (Experiment 1: hippocampus = 15.08, PHC = 14.63, PRC = 15.05; Experiment 2: hippocampus = 16.47, PHC = 16.01, PRC = 15.97; see Table 1). While the SNR distributions suggested that the data acquisition quality in each region was similar, SNR does not provide information about the functional relevance of each voxel nor does it match the number of voxels across ROIs. To that end, the 200 voxels with the largest ANOVA F-scores of activation differences between the two stimulus categories were selected from the training data prior to classification analysis. A similar pattern of results emerged as the full-ROI method—paired t-tests revealed no significant differences between the results obtained from full-ROI and voxel selection methods (all $t_{11} < 1.1$, uncorrected $p > 0.3$). In Experiment 1, one sample t-tests revealed significant classification accuracy for faces versus objects in PHC and PRC (mean classification accuracy: PHC = 59.1%, PRC = 64.9%; both $t_{11} > 7$, corrected $p < 0.001$) but not in the hippocampus (mean classification accuracy = 50.2%, $t_{11} = 0.16$, uncorrected $p = 0.88$; see Table 1). Paired t-tests revealed significantly greater classification accuracy for faces versus objects in PHC and PRC than the hippocampus (both $t_{11} > 5$, corrected $p < 0.005$). In contrast to the full-ROI method, a paired t-test revealed marginally significantly greater classification accuracy for faces versus objects in PRC than PHC ($t_{11} = 3.14$, corrected $p = 0.056$). In Experiment 2, one sample t-tests revealed significant classification accuracy for faces versus scenes in PHC and PRC (mean classification accuracy: PHC=78.1%; PRC=61.6%; both $t_{11} > 8.5$, corrected $p < 0.001$) with little evidence for category discrimination in the hippocampus (mean classification accuracy=52.0%, $t_{11} = 2.17$, uncorrected $p = 0.053$, corrected $p 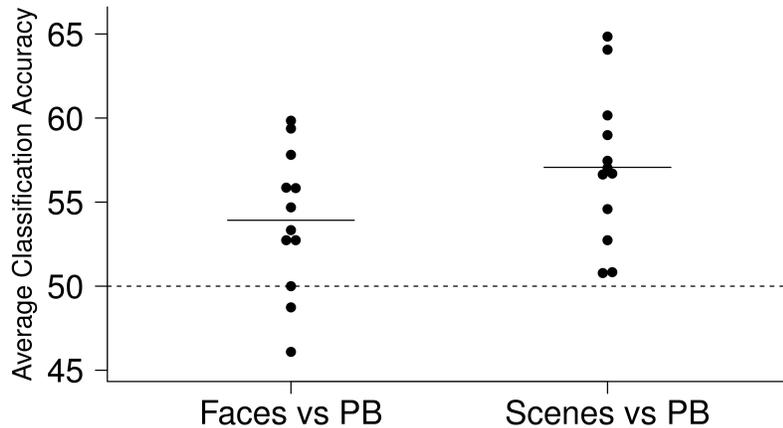= 0.32$; see Table 1). Paired t-tests revealed significantly greater classification accuracy for faces versus scenes in PHC and PRC than the hippocampus (both $t_{11} > 5.5$, corrected $p < 0.005$) and in PHC than PRC ($t_{11} = 6.58$, corrected $p < 0.001$).

To rule out the possibility that the hippocampus was representationally categorical but

	Region	SNR	Standard	Voxel selection	Run averaged
Experiment 1: Faces vs objects	Hippocampus	15.08	51.4	50.2	52.1
	PHC	14.63	61.1^a	59.1^a	82.0^a
	PRC	15.05	64.9^a	64.9^a	91.4^a
Experiment 2: Faces vs scenes	Hippocampus	16.47	52.7	52.0	56.3
	PHC	16.01	$77.7^{a,b}$	$78.1^{a,b}$	$97.1^{a,b}$
	PRC	15.97	62.3^a	61.6^a	81.1^a

Table 2.1: Results of Control Analyses. Abbreviations: SNR, median signal to noise ratio; a, Significant classification accuracy comparted to chance at $p < 0.001$, corrected and significantly greater classification accuracy than the hippocampus at $p < 0.005$, corrected; b, Significantly greater classification accuracy than PRC at $p < 0.001$, corrected.

was not showing significant classification accuracy due to noisier individual trial patterns of activity compared to MTLC, we generated run averaged patterns of activity by averaging all of the individual trial patterns of activity within a category for each run. The results obtained from this analysis can be thought of as being similar those obtained from a block design. A similar pattern of results was observed, with significantly greater classification accuracy in the MTLC using run averaged patterns of activity compared to individual trial patterns of activity for faces versus objects (both $t_{11} > 8$, corrected $p < 0.001$; but not in the hippocampus, $t_{11} = 0.16$, uncorrected $p = 0.87$) and for faces versus scenes (both $t_{11} > 8$, corrected $p < 0.001$; but not in the hippocampus, $t_{11} = 1.26$, uncorrected $p = 0.23$). In Experiment 1, one sample t-tests revealed significant classification accuracy of faces versus objects in PHC and PRC (mean classification accuracy: PHC = 82.0%, PRC = 91.4%; both $t_{11} > 8.5$, corrected $p < 0.001$) but not in the hippocampus (mean classification accuracy = 52.1.%, $t_{11} = 0.47$, uncorrected $p = 0.65$; see Table 1). Paired t-tests revealed significantly greater classification accuracy for faces versus objects in PHC and PRC than the hippocampus (both $t_{11} > 4.54$, corrected $p < 0.005$). In contrast to the full-ROI method using individual trial patterns of activity—but supporting the voxel selection method using individual trial patterns of activity—a paired t-test revealed marginally significantly greater classification accuracy for faces versus objects in PRC than PHC ($t_{11} = 3.17$, corrected $p = 0.053$). In Experiment 2, one sample t-tests revealed significant classification accuracy

for faces versus scenes in PHC and PRC (mean classification accuracy PHC = 97.1%, PRC = 81.1%; both $t_{11} > 15$, corrected $p < 0.001$) with little evidence for category discrimination in the hippocampus (mean classification accuracy = 56.3%, $t_{11} = 2.13$, uncorrected $p = 0.06$, corrected $p = 0.34$; see Table 1). Paired t-tests revealed significantly greater classification accuracy for faces versus scenes in PHC and PRC than the hippocampus (both $t_{11} > 6.5$, corrected $p < 0.001$) and significantly greater classification accuracy in PHC than PRC ($t_{11} = 7.11$, corrected $p < 0.001$).

To investigate whether there were differences between anterior and posterior hippocampus that might have been obscured by the overall poor classification performance, separate analyses were run on each region. The most posterior slice of anterior hippocampus was labeled as the most posterior coronal slice in which the uncal apex was visible (Poppenk and Moscovitch, 2011). A one-sample t-test revealed that classification accuracy for faces versus objects was not significantly different than chance in either region (average classification accuracy: anterior hippocampus = 49.3%, $t_{11} = 0.85$, uncorrected $p = 0.42$; posterior hippocampus = 51.4%, $t_{11} = 1.26$, uncorrected $p = 0.23$) and a paired t-test revealed that classification accuracy in posterior and anterior hippocampus were not significantly different ($t_{11} = 1.42$, uncorrected $p = 0.18$). A one-sample t-test revealed that classification accuracy for faces versus scenes was not significantly different than chance in either region (average classification accuracy: anterior hippocampus = 51.8%, $t_{11} = 1.08$, uncorrected $p = 0.30$, posterior hippocampus = 51.5%, $t_{11} = 1.05$, uncorrected $p = 0.31$) and a paired t-test revealed that classification accuracy in posterior and anterior hippocampus were not significantly different ($t_{11} = 0.13$, uncorrected $p = 0.90$).

To test the hypothesis that RSC/PCC is representationally similar to PHC, we performed classification analysis as well as a variant of informational connectivity. In Experiment 1, a one-sample t-test revealed significant classification accuracy for faces versus objects in RSC/PCC (mean classification accuracy = 55.9%, $t_{11} > 4$, $p < 0.005$, see Figure 2.4A).
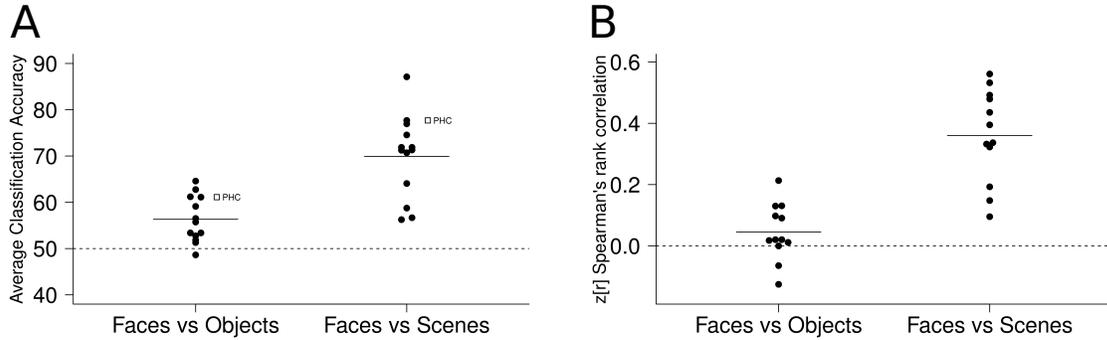
Figure 2.4: (A) Linear SVM classification results in RSC/PCC revealed a similar pattern of results to PHC: significant classification accuracy for both faces versus objects and faces versus scenes, with greater accuracy for faces versus scenes. The white squares represent mean classification results in PHC. (B) Significant informational connectivity between RSC/PCC and PHC was observed in Experiment 2 (faces versus scenes) but not Experiment 1 (faces versus objects).

In Experiment 2, a one-sample t-test revealed significant classification accuracy for faces versus scenes in RSC/PCC (mean classification accuracy $= 69.3\%$, $t_{11} > 7$, $p < 0.001$, see Figure 2.4A). Similar to the results in PHC, a two-tailed Welchs two sample t-test revealed significantly greater classification accuracy for faces versus scenes than faces versus objects ($t > 4$, $p < 0.001$). We used trial-by-trial variability in the SVM decision value to investigate the correlation between multivariate pattern discriminability in PHC and RSC/PCC (see Methods). Two-tailed one sample t-tests revealed a significant Fisher's r to z transformed Spearman's rank correlation between the multivariate pattern discriminability in these regions for faces versus scenes (mean $= 0.36$, $t_{11} = 8.2$, $p < 0.001$, see Figure 2.4B) but not for faces versus objects (mean $= 0.045$, $t_{11} = 1.7$, $p = 0.12$). A two-tailed Welchs two sample t-test revealed significantly greater Fisher's r to z transformed Spearman's rank correlation between PHC and RSC/PCC for faces versus scenes than for faces versus objects ($t = 6.13$, $p < 0.001$).

To further validate our informational connectivity analysis, we conducted two control analyses. First, we simulated data and confirmed that our implementation of informational connectivity is sensitive to continuous distance from the decision boundary. In fact, it is

only sensitive to binary face/non-face information to the degree that the two regions have similar trial-by-trial correct/incorrect decisions. Our simulations revealed that regions that exhibited roughly 75% classification accuracy (similar to our results for faces versus scenes in PHC and RSC/PCC) would not necessarily exhibit significant informational connectivity; in fact, the mean z[r] Spearman's rank correlation of 10,000 simulations was 0.0003, with an approximately Gaussian distribution. Next, we ran single-subject-level permutation analysis, in which we permuted the RSC/PCC vector 10,000 times and calculated the Spearman's rank correlation between each of the permuted vectors and the intact PHC vector. We then calculated the non-parametric two-tailed p-value using the following equation (see Ernst, 2004):

$$p = \frac{1 + \sum_{i=1}^{10,000} I(|t_i - \bar{t}| \geq |t^* - \bar{t}|)}{1 + 10,000} \tag{2.1}$$

where $I(\cdot)$ is the indicator function which sets the value to 1 if the statement is true and to 0 otherwise, t_i represents the ith value of the permutation vector, $\bar{t}$ represents the mean value of the permutation vector, and t^* represents the empirical mean. In other words, we calculated the probability that a null value was at least as far (in both directions) from the mean of the null distribution as the empirical value. This procedure was run on each participant independently. For faces versus objects, Spearman's rank correlation was significantly ($p < 0.05$, uncorrected) greater than the mean permutation value for 3 of 12 participants and was significantly less than the mean permutation value for 1 of 12 participants (all other $p > 0.05$, uncorrected). For faces versus scenes, Spearman's rank correlation was significantly greater than the mean permutation value for 11 of 12 participants, and the 12th participant's p-value was 0.065. Taken together, the simulation and permutation analyses suggest that informational connectivity provides further information about the functional relatedness of PHC and RSC/PCC. Furthermore, this information cannot be explained by a simple binary difference in the level of classification accuracy.

2.4 Discussion

Using biologically-motivated computational models of the MTL, Norman (2010) suggested that the MTLC can compute summed similarity while the hippocampus cannot—a pattern of results caused by differences in the degree of pattern separation in the MTLC and the hippocampus. To test the summed similarity hypothesis, we used high resolution fMRI and MVPA to study whether the MTLC was more representationally categorical than the hippocampus. If the MTLC was more representationally categorical than the hippocampus, it would provide support for the hypothesis that the MTLC is more functionally specialized for computing summed similarity than the hippocampus. Consistent with the hypothesis, linear SVM analysis revealed significant classification accuracy of faces versus objects and faces versus scenes in PHC and PRC with limited evidence at best for category discrimination in the hippocampus. Additionally, greater classification accuracy was observed in PHC and PRC than the hippocampus. These results support the hypothesis that MTLC can compute summed similarity because stimuli within a category (i.e., similar stimuli) elicited similar patterns of activity (with distinct patterns of activity across image categories), while suggesting that the hippocampus cannot.

We utilized voxel selection to control for differing numbers of voxels across ROIs, as well as differing numbers of noisy voxels across ROIs. A similar pattern of results emerged, suggesting that the results were not driven by differing numbers of voxels between MTLC and hippocampus. Further, SNR was similar between ROIs, suggesting that differences in classification accuracy between the MTLC and hippocampus were not driven by differences in the quality of data acquisition. We used run averaged patterns of activity to control for differences in the noisiness of single trial patterns of activity across ROIs. Once again, a similar pattern of results emerged, suggesting that the results were not driven by noisier single trial patterns of activity within the hippocampus. Taken together our results provide novel evidence that the hippocampus is more representationally agnostic than MTLC, a

pattern of results that is not completely driven by non-responsiveness, number of voxels, or noisier individual trial patterns of activity. These results are concordant with the hypothesis that the hippocampus is involved in pattern separation (McClelland et al., 1995; O'Reilly and Munakata, 2000; O'Reilly and Norman, 2002; Norman and O'Reilly, 2003; Norman, 2010) as well as theories that suggest that the hippocampus plays a domain general role in memory formation (Eichenbaum and Cohen, 2001; Davachi, 2006; Azab et al., 2014).

Our results are in agreement with recent studies using MVPA to investigate representations in PHC, PRC, and the hippocampus (Diana et al., 2008; Liang et al., 2013a; LaRocque et al., 2013), which all found greater category representation in the MTLC than the hippocampus. As mentioned in the Introduction, we hypothesize that a core function of the hippocampus is to perform pattern separation and amplify the dissimilarity across representations, which would result in more representationally agnostic patterns of activity than the MTLC. We expect that under certain conditions, above chance classification accuracy for categories could be observed in the hippocampus; however, we posit that even under these conditions the hippocampus would be more representationally agnostic than the MTLC. In fact, Liang et al. (2013a) found evidence that posterior hippocampus was representationally categorical for scenes, while anterior hippocampus was more representationally agnostic. We did not observe significant classification accuracy in posterior (or anterior) hippocampus for faces versus scenes (or faces versus objects), suggesting that our hippocampal results were not driven by differences between these subregions. Overall, however, the results of Liang et al. (2013a) support our conclusions because they found that the hippocampus—including posterior hippocampus—was more representationally agnostic than the MTLC.

Two of the previously mentioned studies also found differences in category representation between PHC and PRC (Liang et al., 2013a; LaRocque et al., 2013). Consistent with previous results, in Experiment 2, greater classification accuracy for faces versus scenes was observed in PHC than PRC. Comparing the results from Experiment 1 and 2, PHC showed greater

classification accuracy for faces versus scenes than faces versus objects. Taken together, these results suggest that PHC may be preferentially tuned for representation of scenes; however, the significant classification of faces versus objects in the PHC suggests that it is not involved solely in representation of scenes. We found marginally significantly greater classification accuracy in PRC than PHC for faces versus objects when using voxel selection of individual patterns of activity and when using run averaged patterns of activity, suggesting that PRC may be relatively specialized for representing object or face stimuli. Taken together, our findings support the binding of item and context model (Eichenbaum et al., 2007; Diana et al., 2007), which posits that PHC is relatively specialized for representing contexts (including spatial and non-spatial contextual information, as well as scenes, based on anatomical input from neocortical structures in the where pathway) and PRC is relatively specialized for representing items (e.g., objects, based on anatomical input from neocortical structures in the what pathway). Furthermore, our findings add to studies implicating PHC's role in scene representation (Litman et al., 2009; Liang et al., 2013a; see Davachi, 2006, for a review of studies using univariate analysis of fMRI to investigate domain specificity in PHC, PRC, and the hippocampus).

Recent theories suggest that RSC should be included as part of a larger MTL-cortical memory network (Vann et al., 2009; Ranganath and Ritchey, 2012). Expanding on the binding of item and context model, Ranganath and Ritchey (2012) hypothesized that PHC, RSC, and PCC are part of a posterior memory network involved in contextual representation. Consistent with our hypothesis that PHC and RSC/PCC are representationally similar, we observed a similar pattern of results in RSC/PCC to PHC: significant classification of faces versus objects and faces versus scenes, with significantly greater classification for faces versus scenes than faces versus objects. Similar to our conclusions of PHC, these findings suggest that RSC/PCC might be preferentially tuned for scene or spatial representation; however, the significant classification of faces versus objects in RSC/PCC suggests that it is not tuned solely for the representation of scenes or spatial information. The fact that RSC/PCC

showed significant classification in both experiments further highlights the distinctness of the representationally agnostic results observed in the hippocampus.

A variant of informational connectivity (Coutanche and Thompson-Schill, 2013) was used to investigate the correlation between trial-by-trial multivariate pattern discriminability in PHC and RSC/PCC. As discussed previously (Coutanche and Thompson-Schill, 2013), informational connectivity provides novel insight into the functional synchrony between brain regions by measuring the relatedness in multivariate pattern discriminability over time. Informational connectivity is sensitive both to the trial-by-trial decision (e.g., face or non-face) as well as the continuous distance from the decision boundary. Importantly, similar classification accuracy between regions does not ensure significant informational connectivity, therefore informational connectivity can provide further information about whether brain regions are representationally similar. Consistent with the hypothesis that PHC and RSC/PCC are functionally related and involved in representing scenes or spatial information, we observed a significant correlation between the multivariate pattern discriminability on a trial-by-trial basis in these regions for an experiment with face and scene stimuli but not an experiment with face and object stimuli. These findings suggest that there may be a stimulus-dependent modulation in the correlation between multivariate pattern discriminability across these regions. Further, these results suggest that PHC and RSC/PCC contain similar trial-by-trial information in the task with faces and scenes.

The hippocampus showed little evidence for classification accuracy for faces versus objects and faces versus scenes, suggesting that the hippocampus is not representationally categorical for these categories. As suggested by Diana et al., 2008, the lack of evidence for category representation in the hippocampus does not rule out the possibility that it is representationally categorical at a more fine-grained scale than our fMRI resolution. However, the significant classification accuracy for images (faces and scenes) compared to perceptual baseline suggests that the hippocampus responded to the images. Furthermore, in contrast to the

relative lack of evidence of category representation in the human hippocampus (Diana et al., 2008; LaRocque et al., 2013; but see Liang et al., 2013a), recent evidence using linear SVM analysis showed successful classification between distinct locations in a virtual environment (Hassabis et al., 2009), between memories of highly overlapping video clips (Chadwick et al., 2010, 2011), and between highly similar scene images (Bonnici et al., 2012). Additionally, these studies observed significantly higher classification accuracy in the hippocampus than the MTLC. These results provide evidence for the role of the human hippocampus in pattern separation, thus suggesting that the hippocampus may exhibit unique patterns of activity in response to each individual stimulus, thus resulting in a lack of category representation (cf. Chadwick et al., 2012). Notably, the aforementioned studies used 1.5 mm isotropic voxels, thus suggesting that linear SVM classification analysis is sufficient to reveal patterns of activity within the hippocampus using the same resolution as we did in our experiments. These findings address the concern that the hippocampus contains patterns of activity that are present at a level of resolution that is too fine-grained to observe in our current data set.

The behavioral task in our experiment was an incidental memory task, which raises the concern that participants were not required to memorize the stimuli. Theoretical accounts (e.g., Martin, 1999; O'Reilly and Rudy, 2001) and empirical evidence (e.g., Rugg et al., 1997; Otten et al., 2001; Stark and Okado, 2003) have suggested that the hippocampus encodes information regardless of task demands. Furthermore, as previously mentioned, the significant classification accuracy for images compared to the perceptual baseline task in the hippocampus suggests that it was responding to the images. LaRocque et al. (2013) used an incidental encoding task to provide support for behavioral relevance of the differences in representations in the MTL cortex and hippocampus. Their results revealed a positive relationship between subsequent memory and within-category similarity in PHC and PRC and a negative relationship between subsequent memory and within-category similarity in the hippocampus. Taken together, it appears that there may be a double dissociation between the MTLC and the hippocampus, such that the MTLC is specialized for computing summed similarity

and the hippocampus is specialized for pattern separation, thus supporting computational theories of the MTL (Norman, 2010).

Chapter 3

What's in a context? The influence of low-level stimulus features on representations in the human medial temporal lobe and retrosplenial cortex.

Our ability to remember the events of our lives critically relies upon the formation of associations among the "what", "where", and "when" components of the event. The context-guided object association task (e.g., McKenzie et al., 2014) has provided a useful task to investigate the representation of events within subregions of the rodent medial temporal lobe, regions which are known to be necessary for declarative memory. We created two human versions of the context-guided object association task for use in functional magnetic resonance imaging experiments. In the first experiment, we used distinct items and distinct contexts, similar to the rodent task. The results of our first experiment largely replicate previous work in the rodent and extend their findings to human participants and to other brain regions. Specifically, the results provide evidence which is consistent with the representation of context in the hippocampus, parahippocampal cortex, and retrosplenial cortex. The results also pro-

vide evidence which is consistent with the conjunctive representation of items-in-context and with a relationship to behavior in retrosplenial cortex. However, we also found similar relationships in primary visual cortex, suggesting that low-level differences between the stimuli could be influencing representations in the MTL and retrosplenial cortex. In the second experiment, we used stimulus filtering and model testing to investigate whether the results from our first experiment would maintain in the absence of low-level differences between stimuli. The results of our second experiment provide little evidence for invariant context representation in the hippocampus, parahippocampal cortex, and retrosplenial cortex, suggesting that the results from our first experiment were influenced by the low-level visual differences between stimuli. In contrast, our results provide novel evidence for the role of perirhinal cortex in invariant object representation. The results from our experiments provide insight into the types of experiments that should be used to investigate the representation of items and contexts; accordingly, we will conclude by discussing possible experimental designs for future studies.

3.1 Introduction

Following the discovery that removal of structures within the human medial temporal lobe (MTL) causes amnesia (Scoville and Milner, 1957), decades of research have focused on elucidating the contributions of subregions of the MTL to declarative memory. Separately, the discovery of place cells in the rodent hippocampus (O'Keefe and Dostrovsky, 1971) initiated a line of research which has focused on the contributions of MTL subregions to spatial processing (for reviews on merging these frameworks see: Eichenbaum and Cohen, 2014; Schiller et al., 2015). While there is still debate over the precise nature of the division of labor within the MTL, there is consensus that the hippocampus sits at the apex of a cortical circuit that allows unparalleled it access to the what, where, and when components of events and that it,

in some way, is crucial for binding or associating these aspects (Mishkin et al., 1997; Cohen et al., 1999; Lavenex and Amaral, 2000; Davachi, 2006; Diana et al., 2007; Eichenbaum et al., 2007; Wixted and Squire, 2011; Ranganath and Ritchey, 2012). Converging evidence suggests that adjacent MTL cortical regions are also necessary for declarative memory, albeit perhaps in a more domain-specific manner (Mishkin et al., 1997; Davachi, 2006; Diana et al., 2007; Eichenbaum et al., 2007; Wixted and Squire, 2011; Ranganath and Ritchey, 2012). Many theoretical accounts note a role for parahippocampal cortex (PHC; called postrhinal cortex in rodents) and medial entorhinal cortex in memory for and in the representation of context and spatial information (Burwell, 2000; Davachi, 2006; Knierim et al., 2006; Diana et al., 2007; Eichenbaum et al., 2007; Wixted and Squire, 2011; Ranganath and Ritchey, 2012). On the other hand, perirhinal cortex (PRC) and lateral entorhinal cortex are hypothesized to be involved in memory for and in the representation of objects.

In a series of groundbreaking studies, Eichenbaum and colleagues developed a context-guided object association task to study associative memory in rodents. Briefly, animals learn item-reward contingencies that differ based on the context (operationally defined as visually and tactilely distinct chambers). The results of their experiments have firmly established the involvement of the hippocampus (Rajji et al., 2006; Komorowski et al., 2009, 2013; Navawongse and Eichenbaum, 2013; Tort et al., 2013; McKenzie et al., 2014), medial temporal lobe cortical regions (Keene et al., 2016), and orbitofrontal cortex (Farovik et al., 2015) in different aspects of task performance (for review see: McKenzie et al., 2015). Impaired context-guided object association learning has been shown in rats with hippocampal lesions (Komorowski et al., 2013) and in mice with impaired NMDA receptor function (Rajji et al., 2006), thus establishing a necessary role for the hippocampus in task performance. Neurophysiological investigations (Komorowski et al., 2009, 2013; Navawongse and Eichenbaum, 2013; Tort et al., 2013; McKenzie et al., 2014) have indicated that the hippocampus represents not only spatial aspects of the task (i.e., context, location) but also non-spatial components of the task (i.e., valence, item). Moreover, McKenzie et al. (2014) found that the hippocam-

pus contains hierarchically-organized representations, such that patterns of activity differ in a graded manner based on differences in the stimuli that comprise an event (ranked from most dissimilar to most similar): 1) events that take place in different contexts, 2) events that take place at different locations, 3) events with a different valence (i.e., rewarded or unrewarded), 4) events that contain different items, 5) events with all of the same stimuli. Collectively, these results strongly support the notion that the hippocampus has access to multiple features of task performance (e.g., what, where), corroborating relational theories of hippocampal processing (e.g., Eichenbaum et al., 2007).

A recent study that used the context-guided object association task suggested that medial entorhinal cortex represents not only spatial aspects but also non-spatial aspects of the task, similar to the results from the hippocampus (Keene et al., 2016). Additionally, their results suggested that PRC and lateral entorhinal cortex represent not only item information but also represent spatial aspects of the task. Interestingly, however, representational similarity analysis demonstrated that patterns of activity in medial entorhinal cortex were more dissimilar in response to events that took place in different locations than events that contained different items. Conversely, patterns of activity in PRC and lateral entorhinal cortex were more dissimilar in response to events that contained different items than events that took place in different locations. Taken together, Keene et al. (2016) argue that their results challenge theories that propose a simple dissociation between medial entorhinal cortex and lateral entorhinal cortex and PRC but that their results suggest that there are representational differences between these regions (cf. Knierim et al., 2013).

In the present report, we developed a human version of the context-guided object association task for functional magnetic resonance imaging (fMRI). In addition to investigating subregions of the MTL, we sought to investigate other regions that have been hypothesized to be involved in memory and context processing. For example, human neuroimaging studies (Bar and Aminoff, 2003; Park and Chun, 2009; Walther et al., 2009; Auger and Maguire, 2013;

Auger et al., 2015; Wing et al., 2015) and lesion and neurophysiological studies in the rodent (Ennaceur et al., 1997; Vann and Aggleton, 2002; Parron and Save, 2004; Chen et al., 1994; Cho and Sharp, 2001; Alexander and Nitz, 2015) suggest that retrosplenial cortex (RSC), a subregion of the posterior cingulate cortex (PCC), is involved in processing scenes, context, and spatial information. Additionally, RSC is reciprocally connected to the hippocampus, PHC, and anterior thalamic nuclei, which are critically important for declarative memory (Vann et al., 2009; Aggleton, 2010). Indeed, RSC lesions are accompanied by "retrosplenial amnesia" (Valenstein et al., 1987) and recent accounts postulate that RSC should be thought of as a part of the MTL-declarative memory system (Ranganath and Ritchey, 2012; Vann et al., 2009; Aggleton, 2010). Furthermore, a recent model suggests that RSC is critical for the formation of stimulus-stimulus associations (Bucci and Robinson, 2014), a function traditionally ascribed to the hippocampus. Thus, a more complete understanding of the contribution of structures beyond the MTL, especially RSC, to declarative memory is of keen interest. We also tested the hypothesis that, similar to findings in medial entorhinal cortex (Keene et al., 2016), representations in RSC would not only carry information about context but also about other aspects of the task.

In the context-guided object association task, the contexts and items are composed of distinct elements (i.e., different visual, olfactory, and tactile cues). An unaddressed question is the extent to which the neural representations are influenced by the low-level differences between stimuli. The use of distinct contexts and items allows animals (e.g., rats, humans) to rapidly discriminate between them (cf. Bulkin et al., 2016); however, as we discovered in the course of the present report, it is virtually inevitable that representational differences will also be present in the relevant primary sensory regions. We propose that the cognitive representation of a context should theoretically be stable across different versions of the same context so long as the context signals a reliable behavioral outcome (e.g., Context A + Item X = Response 1). Given the previous finding that patterns of activity in the hippocampus and MTL cortex were maximally dissimilar for events that took place in a different context (McKenzie et al., 2014;

Keene et al., 2016), the second aim of this study was to examine whether representations of context in the MTL and RSC are dependent on low-level differences between contexts or whether such representations would exhibit invariance.

In Experiment 1, we used distinct stimuli for our contexts (time lapse videos) and objects (images of objects). In Experiment 2, we used a computational approach to match the low-level visual features of our stimulus set in order to test for invariant context and object representation. Similar to previous findings in the rodent (McKenzie et al., 2014; Keene et al., 2016), the results of Experiment 1 provide evidence for the representation of context within subregions of the MTL. Additionally, our results suggest that RSC carries context and conjunctive item-in-context information and such representations are related to behavioral performance. We also observed evidence for context and conjunctive item-in-context information in other scene processing regions, parahippocampal place area and retrosplenial complex (e.g., Epstein and Kanwisher, 1998; Epstein et al., 2007; Julian et al., 2012; Vass and Epstein, 2013, 2016; Marchette et al., 2014, 2015). However, we observed similar relationships in primary visual cortex. In contrast, the results of Experiment 2 provide little evidence for invariant context representation within the same regions, suggesting that the results from Experiment 1 were influenced by low-level stimulus features. Moreover, these results highlight the potential for low-level stimulus features to masquerade as contextual information, thus providing clear boundary conditions for the investigation of context representation. These results also raise interesting questions about how to define a context. For example, while the context-guided object association task has defined contexts by physically distinct chambers, other studies have advocated for the use of more psychological or behavioral-based contexts (Smith and Mizumori, 2006a; Smith and Bulkin, 2014). In contrast to the relatively null results of invariant context representation, the results of Experiment 2 provide novel evidence that PRC is involved in the invariant representation of objects. We conclude by providing suggestions for future experiments that aim to investigate context and object representation in the absence of low-level confounds.

3.2 Materials and Methods

3.2.1 Participants

Thirty-five participants were recruited from the community at the University of California, Irvine. Participants were between 18 and 31 years of age, were right handed, and screened negative for neurological and psychiatric disease. Five participants were excluded due to excessive motion. Twenty participants were included in the analysis in Experiment 1 (10 females, 10 males) and ten in Experiment 2 (5 females, 5 males). Participants consented to the procedures in accordance with the Institutional Review Board of the University of California, Irvine, and received monetary compensation for their participation.

3.2.2 Stimuli

Experiment 1: Distinct stimulus set

In Experiment 1, the stimulus set consisted of two time-lapse videos (clips from Timestorm Films: `https://vimeo.com/93003441`; played full screen; monitor resolution: 1280 × 1024 pixels) and two object pairs (150 × 150 pixels; Fig. 3.1A).

Experiment 2: Low-level image matching

In Experiment 2, the two contexts consisted of grayscale images (600 × 600 pixels) of Saint Peter's Basilica and the U.S. Capitol Building and the objects consisted of grayscale images (256 × 256 pixels) of car and house keys. We used a combined approach of image manipulation and model testing to diminish the presence of category information from the low-level visual features. First, we used the SHINE toolbox (Willenbockel et al., 2010) to
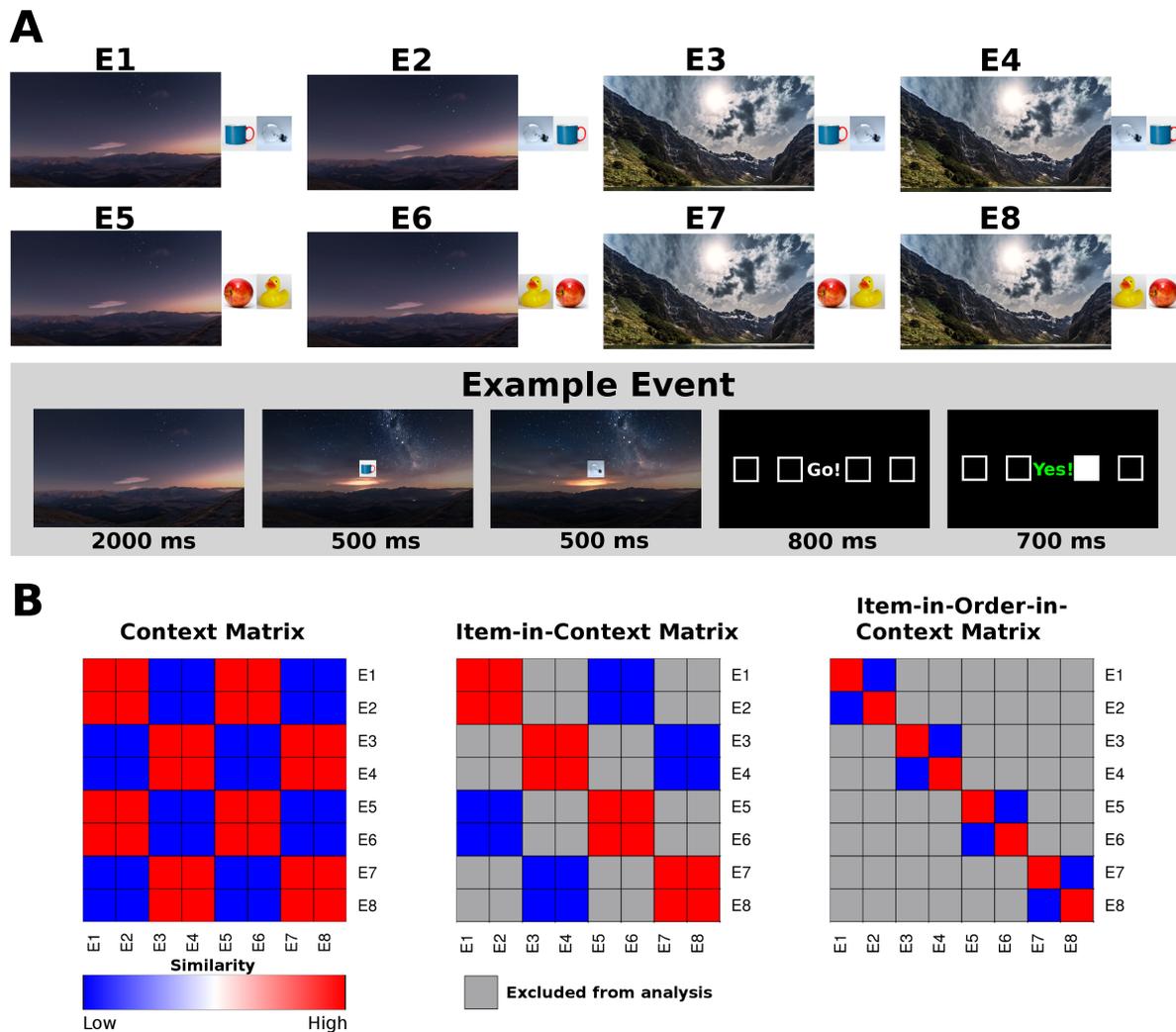
Figure 3.1: Experiment 1 stimuli, event design, and model matrices. A) The task stimuli and an example event. The stimulus set consisted of two time-lapse videos (clips from Timestorm Films: https://vimeo.com/93003441) and two object pairs. Each event began with a 2000 ms presentation of a time-lapse video (depicted by the scene), then an object was displayed at the center of the video for 500 ms (depicted by the object on the left), which was then replaced by a second object which was displayed for 500 ms (depicted by the object on the right). Participants learned event-location associations the day prior to scanning, and were tested on the associations during scanning. B) Model matrices for our representational similarity analysis.

equate luminance histograms across all of the scene stimuli and across all of the object stimuli (Fig. 3.2A). Second, we used a modeling approach to select images that were devoid of low-level category features.

For our scene images, similar to Marchette et al. (2015), we used pixel-wise correlation, the GIST computational model (Oliva and Torralba, 2001), and the HMAX computational model (two variants, one which used all images from the Fifteen Scene Categories dataset [Lazebnik et al., 2006] as prototypes and one which used a superset of our scene images as prototypes; we used the model from: Theriault et al., 2011). Additionally, similar to Kriegeskorte et al. (2008a), we used two models of V1 (one which included both simple and complex cells from HMAX and another that included only complex cells; Theriault et al., 2011), low-pass pixel-wise correlation (low frequency image features), high-pass pixel-wise correlation (high frequency image features), and Radon transform. We iteratively looped over a superset of our scene images and selected images for which all nine models showed no sign of a relationship between the scene images and the context matrix for both the selected stimulus set (40×40 matrix with 780 unique entries) and across the odd/even split (20×20 matrix with 400 unique entries; $-0.012 <$ Spearman's rank correlation < 0.011, all p's > 0.77). As a final control, we simulated an object being presented on top of each scene image by placing a black square (256×256 pixels) at the center of the image; importantly, similar results were obtained using this method.

For the object images, we used the same nine models as well as binary-silhouette correlation (similar to Kriegeskorte et al., 2008a). Similar to the analysis of the scene images, we used two versions of the HMAX model (one which used all images from the 256 Object Categories dataset [Griffin et al., 2007] as prototypes and one which used a superset of our object images as prototypes; Theriault et al., 2011). We iteratively looped over a superset of our object images and selected images for which all ten models showed no sign of a relationship to the object matrix for both the selected stimulus set (40×40 matrix with 780 unique entries)

and across the odd/even split (20×20 matrix with 400 unique entries; $-0.016 \leq$ Spearman's rank correlation ≤ 0.014, all p's > 0.74).

3.2.3   Behavioral task

Experiment 1: Pre-scan training task

Participants were trained on an associative memory task one day prior to their scan session. The behavioral task was created using custom-written code and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Each event began with a 2000 ms presentation of a time-lapse video. An object was then displayed at the center of the video for 500 ms, which was replaced by a second object which was displayed for 500 ms, resulting in a 3 second event duration. One difference from the rodent task was that rather than using reward per se, participants learned stimulus-location associations through trial and error learning (Fig. 3.1A), similar to previous experiments in our laboratory (Law et al., 2005). In Experiment 1, the correct location was unique based on: 1) the video that was displayed, 2) the objects that were displayed, and 3) the order in which the objects were presented. Thus, the task required participants to make distinct responses for events that contain overlapping features. Participants responded using the 4 fingers on their right hand, and the event-location contingency was balanced across the two contexts. We used a response window of 800 ms, followed by 700 ms of feedback ("Yes!", "No!", or "?" = no response). The interstimulus interval consisted of a 400 ms fixation cross, a 700 ms arrow presentation (to which participants were instructed to indicate via button press whether it was pointing to the left or the right, which served as a non-mnemonic component of the interstimulus interval), and a 400 ms fixation cross, resulting in a trial length of 6 seconds (event duration = 3 seconds, response and feedback = 1.5 seconds, and interstimulus interval = 1.5 seconds). We also included self-paced perceptual baseline trials (5.6 second blocks followed by a 400 ms fixation cross)

in which participants were instructed to indicate as quickly and as accurately as possible which of four static noise boxes was the brightest (Law et al., 2005). The brightness of the target box was continuously titrated to maintain performance between 40 and 60% correct.

At the beginning of the experiment, two events were presented. After a participant learned an event-location association to criterion—correct responses on 5 out of the last 6 responses—a new event was added to the unlearned queue, and the learned event was moved into the learned queue. Events in the learned queue were presented with p = 0.3, while events in the unlearned queue were presented with p = 0.7. Thus, participants continued to be tested on "learned" event-location pairs. We counterbalanced the order in which events were added to the unlearned queue. The session terminated after participants learned all 8 events to criterion.

Experiment 2: Pre-scan training tasks

To ensure that participants could readily discriminate between the the "context" images that we used in the associative learning task, they were first trained on a category discrimination task for the 40 scene images. On each trial an image of either Saint Peter's Basilica or the U.S. Capitol Building was displayed for 1000 ms. Then, the image was removed and two boxes were displayed with text labels above each box ("St. Peter's Basilica" and "U.S. Capitol Building"; the left/right assignment of the text labels was random on each trial). We used

Figure 3.2: Experiment 2 stimuli, event design, and model matrices. A) Task stimuli and an example event. The stimulus set consisted of 20 images of Saint Peter's Basilica, 20 images of the U.S. Capitol Building, 20 images of car keys, and 20 images of house keys. The odd numbered rows were used in odd runs of the task while the even numbered rows were used in even runs of the task. Each event began with a 2000 ms presentation of a context image, then an object was displayed at the center of the scene image of 500 ms. B) Model matrices for the representational similarity analysis. E1 = St. Peter's Basilica + Car Key; E2 = St. Peter's Basilica + House Key; E3 = U.S. Capitol Building + Car Key; E4 = U.S. Capitol Building + House Key.

A

Context Images

St. Peter's Basilica

U.S Capitol Building

Object Images

Car Keys

House Keys

Example Event

| 2000 ms | 500 ms | 1300 ms | 700 ms |

B

Context Matrix

E1
E2
E3
E4

E1 E2 E3 E4

Object Matrix

E1
E2
E3
E4

E1 E2 E3 E4
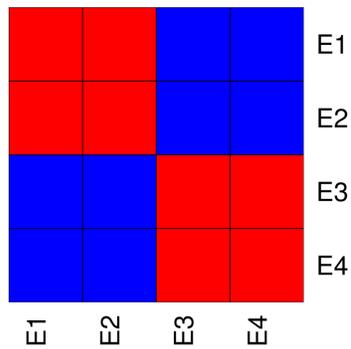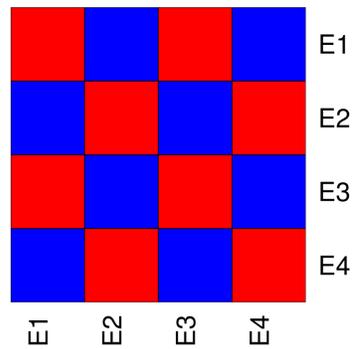
a 1300 ms response window, followed by 700 ms of feedback. The task terminated after participants learned each of the 40 scene images to criterion (correct responses on 5 out of the last 6 responses). In contrast to the associative memory task, only unlearned items were presented. Next, to ensure that participants could readily discriminate between the "object" images that we used in the associative learning task, participants performed an object discrimination task for the car and house key images. The task was identical to the scene task, except the text labels were "car" and "house."

After participants learned both the scene categories and the object categories to criterion, they learned object-location associations. The event structure was similar to Experiment 1, with the exception that we used only one object per event, resulting in a stimulus duration of 2500 ms. Additionally, events were mapped to two responses (left versus right) and participants used their index finger to respond. Given the 500 ms reduction of stimulus presentation, we extended the response window from 800 ms to 1300 ms. Finally, we reduced the perceptual baseline task to contain two boxes, and performance was titrated to maintain performance between 60 and 70% correct.

Experiment 1: fMRI task

Participants returned for their functional magnetic resonance imaging (fMRI) scan session one day after the training phase. During acquisition of structural scans, participants performed a warm-up phase, in which they were re-exposed to the associative memory task. The warm-up phase was included to attenuate novelty effects during the initial presentations of each event (Law et al., 2005). Once again, participants were initially tested on two event-location pairs. After participants learned an event-location association to criterion—two correct responses in a row—a new event was added to the unlearned queue and the learned event was added to the learned queue. As in the initial learning phase, items in the learned queue were presented with p = 0.3, and the warm-up phase terminated after participants

relearned all 8 events to criterion.

We designed the training and imaging paradigm to be similar to that used by (McKenzie et al., 2014), in which they trained rats to criterion prior to neural recording. Thus, both studies investigated well-learned representations. During functional runs, participants were repeatedly tested on the event-location association task. Functional runs consisted of 4 presentations of each event as well as 5 self-paced perceptual baseline trials. We randomized the order of events within each run, with the exception that every run ended with one perceptual baseline trial to allow the hemodynamic response of the 2nd to last trial of the run to approach baseline prior to run completion. Participants completed 16 runs, resulting in 64 presentations of each event during functional scanning.

Experiment 2: fMRI task

There were a few minor differences between the fMRI task in Experiment 1 and Experiment 2. First, participants were given a reminder session for the stimulus categories (i.e., Saint Peter's Basilica, U.S. Capitol Building, car keys, house keys). Participants viewed a text label of the stimulus category, followed by a one second presentation of every image from the category (1 second presentation, 500 ms interstimulus interval). Participants saw each category two times. Second, functional runs consisted of 5 presentations of each event as well as 6 self-paced perceptual baseline trials. Third, de Bruijn sequences were used for stimulus ordering (Aguirre et al., 2011). A unique sequence was used for each run (randomized across subjects), and we selected sequences that ended with perceptual baseline trials in order to allow the hemodynamic response of the 2nd to last trial of the run to approach baseline prior to run completion. Carry-over sequences, such as de Bruijn sequences, match the number of times that each stimulus precedes every other stimulus, thus controlling for stimulus carry-over effects and theoretically increasing the detection power in between-run pattern analysis (Aguirre, 2007; Aguirre et al., 2011). The length of carry-over sequences

was prohibitively large for Experiment 1, but the reduction of stimulus conditions allowed an entire de Bruijn sequence to be presented within a short run (for an argument for using short runs for pattern analysis see: Coutanche and Thompson-Schill, 2012). Finally, participants completed 12 runs, resulting in 60 presentations of each event during functional scanning.

3.2.4   MRI data acquisition

Data were acquired from a 3.0-T Philips scanner, using a 32 channel sensitivity encoding (SENSE) coil at the Neuroscience Imaging Center at University of California, Irvine. A high-resolution 3D magnetization-prepared rapid gradient echo (MP-RAGE) structural scan (0.75 mm^3) was acquired for each participant. Functional MRI scans consisted of a T2*-weighted echo planar imaging sequence using blood-oxygenation-level-dependent contrast (BOLD; repetition time [TR]=2500 ms, echo time=26 ms, flip angle=70 degrees, 46 slices, 2.5×2.5 mm in plane resolution, 2.3 mm slice thickness with a 0.2 mm gap). Each functional run was padded with an initial 4 "dummy" dynamics, which were immediately discarded to ensure T1 stabilization. In Experiment 1, 90 dynamics were collected per run and 16 functional runs were collected for each participant; however, the 16th run for one participant was not analyzed due to large between-run motion. In Experiment 2, 64 dynamics were collected per run and 12 functional runs were collected for each participant.

3.2.5   fMRI data preprocessing

Data were preprocessed using Analysis of Functional NeuroImages (AFNI; Cox, 1996). Functional MRI data were motion corrected using rigid-body transformation using the function align_epi_anat.py (Saad et al., 2009). Data were quadratically detrended and high pass filtered ($f > 0.01$ Hz), using the 3dBandpass function. To preserve fine-grained information, the data were left unsmoothed. We manually defined the hippocampus, parahippocampal

cortex (PHC), and perirhinal cortex (PRC) on a custom template brain according to previously defined landmarks (for more details see: Law et al., 2005; Huffman and Stark, 2014). We used Advanced Normalization Tools (ANTs; Avants et al., 2008) to warp each individual participant's MP-RAGE structural scan into our custom template space. The inverse warp vectors were used to create masks for the hippocampus, PHC, and PRC within each participant's original space (Huffman and Stark, 2014). Freesurfer's isthmus cingulate label (Desikan et al., 2006) was used to define RSC/PCC. The isthmus cingulate mask contains voxels from RSC (traditionally defined as Brodmann's areas 29 and 30; Vann et al., 2009) and a portion of PCC caudal to RSC. Control regions, left primary motor cortex and bilateral V1, were generated using Freesurfer's precentral gyrus label (Destrieux et al., 2010) and Freesufer's V1 atlas (Hinds et al., 2008), respectively. Masks were resampled to 2.5 mm isotropic (the fMRI grid) and further masked to contain completely-sampled voxels. We used a combined anatomical and functional approach to define parahippocampal place area (PPA) and retrosplenial complex (RS-Complex; Julian et al., 2012). We warped the anatomical masks to each subject's native space and selected the 100 most active voxels (all events versus perceptual baseline) in each anatomical mask in each hemisphere and merged the resultant files to create bilateral masks for each ROI (Marchette et al., 2015; Vass and Epstein, 2016).

3.2.6 Representational similarity analysis

Data were analyzed using AFNI, custom-written code in python and R, and PyMVPA (Hanke et al., 2009) on a GNU/Linux platform using the NeuroDebian package repository (Hanke and Halchenko, 2012). The fMRI data were split in half—odd and even runs—and a block-based general linear model (GLM; AFNI's 3dREMLfit function; block length = 3 seconds—i.e., the length of the event presentations) was used to generate beta values in each voxel. During volume registration, 6 motion parameters were generated (3 translation parameters

and 3 rotation parameters). Framewise displacement is defined as the sum of the absolute value of the difference between each of the 6 motion parameters between successive frames (Power et al., 2012). Frames with framewise displacement exceeding 0.5 mm and 1 frame before and 2 frames after were censored from the analysis (for a similar approach to functional connectivity analysis see: Power et al., 2012). Within each split, the mean pattern of activity across all events was subtracted from each event-specific beta vector (Haxby et al., 2001). Pearson's correlation coefficient was calculated between each event-specific beta vector across the split-halves, resulting in non-symmetrical representational similarity matrices. Hypothesis-driven analysis was conducted by calculating Spearman's rank correlation coefficient between each participant's representational similarity matrix and pre-defined matrices (Figs. 3.1B, 3.2B).

We performed an iterative approach in our correlation analysis between each subject's representational similarity matrix and our model matrices. We chose this approach for two reasons: 1) the model matrices were correlated with each other, thus precluding analysis within a single model, 2) correlation analysis is not sensitive to the magnitude of the values within the similarity matrix (as opposed to comparing whether within-category correlations were numerically larger than between-category correlations), thus allowing us to examine the pattern of similarity regardless of the magnitude of the values (Kriegeskorte et al., 2008a). In Experiment 1, we began with the model matrix on the left side of Fig. 3.1B (i.e., the context matrix) and proceeded rightwards only for matrices that were significantly related to the model matrix in the present step (i.e., we terminated analysis for an ROI when the relationship between the ROI matrix and model matrix failed to reach significance). A previous study (Kriegeskorte et al., 2008b) investigated the cross-species correlation between portions of representational similarity matrices, which is similar to our approach of comparing portions of representational similarity matrices to model matrices. We used Spearman's rank correlation, rather than Pearson's correlation coefficient, because it is better suited for investigating the relationship between ordinal models and representational similarity ma-

trices (Kriegeskorte et al., 2008a). Figures were generated using R (built-in packages and the heatmap.2 package within gplots), and we used GNU Image Manipulation Program (`http://www.gimp.org`) to format our figures.

3.2.7 Permutation analysis

Previous reports have suggested that nonparametric methods are preferable to classical statistical tests for analyzing the significance of the relationship between representational similarity matrices (Kriegeskorte et al., 2008a,b); therefore, we used a two-step permutation method to determine statistical significance (for related approaches to classification analysis see: Chen et al., 2011; Liang et al., 2013b; Stelzer et al., 2013; Etzel, 2015). For each participant, the empirical similarity matrix was randomly shuffled and we calculated Spearman's rank correlation between the resultant matrix and the intact model matrix. The resultant value was then Fisher's r-to-z transformed (z[r]), using the inverse hyperbolic tangent function. We performed this process 10,000 times for each participant to generate null distributions at the subject level. To maintain similarity to the empirical analysis, we used the same permutation of the labels across participants (Etzel, 2015). In the current data set this approach was more conservative than using a different permutation of the labels across participants. The null distributions were averaged across participants to generate a null mean z[r] Spearman's rank correlation. Two-tailed nonparametric p value were calculated using the following equation (Ernst, 2004):

$$p = \frac{1 + \sum_{i=1}^{10,000} I(|t_i - \bar{t}| \geq |t^* - \bar{t}|)}{1 + 10,000} \tag{3.1}$$

where $I(\cdot)$ is the indicator function which sets the value to 1 if the statement is true and to 0 otherwise, t_i represents the ith value of the permutation vector, $\bar{t}$ represents the mean value of the permutation vector and t^* represents the empirical mean. This calculates the

probability that a null value was at least as far (in both directions) from the mean of the null distribution as the empirical value. In order to maximize the degree of similarity to a full permutation approach, 1 is added to both the numerator and the denominator of the equation (i.e., in a full permutation the stimulus labels would be in the correct order exactly once, hence the lowest p-value attainable is 1 divided by the number of combinations). For Experiment 2, the full permutations were tractable, thus p values were calculated using the following equation (Ernst, 2004):

$$p = \frac{\sum_{i=1}^{M} I(|t_i - \bar{t}| \geq |t^* - \bar{t}|)}{M} \tag{3.2}$$

where M is the total number of combinations. In Experiment 2, we performed nonparametric difference tests by subtracting permutation matrices from each other across ROIs. The resultant permutation difference matrices were averaged across participants and significance was assessed using Equation 2, where the t's represent difference values.

3.2.8 Informational correlativity analysis

In Experiment 1, we tested the hypothesis that the hippocampus, PHC, and RSC/PCC contain similar information about context on a trial-by-trial basis, using a variant of "informational connectivity" (Coutanche and Thompson-Schill, 2013; Huffman and Stark, 2014), which we refer to here as informational correlativity. We used an extension of the LS2 procedure (Turner et al., 2012) to obtain individual trial estimates of activity. Briefly, for each trial, we ran a block-based GLM analysis (using AFNI's 3dDeconvolve function) that included an individual trial regressor and 8 event-specific regressors that coded for every other trial. We then performed 512 iterations of this procedure (i.e., 8 events × 64 presentations of each event). Each iteration of the GLM incorporated the same censor vectors as before (i.e., framewise displacement > 0.5 mm); additionally, to mitigate the adverse effect

82

of noisy individual trial estimates from subsequent analysis, we removed trials that had a motion event within approximately 15 seconds of the onset of the trial (the exact duration was variable because the image acquisition was not time-locked to the trial presentation).

To maintain similarity to the initial analysis, we used a split-halves approach in this procedure. Similarly, we subtracted the mean pattern of activity across all events from each event-specific beta vector, separately within each split. We then averaged the patterns of activity within each context (i.e., E1, E2, E5, and E6 were averaged to create an average "context 1" pattern of activity and E3, E4, E7, and E8 were averaged to create an average "context 2" pattern of activity) separately within each split. For each trial, we calculated z[r] Pearson's correlation coefficient between the pattern of activity on that trial and the average context 1 and context 2 patterns of activity from the other split. The value for each trial was set to the correlation to the same context minus the correlation to the other context (Coutanche and Thompson-Schill, 2013). Thus, values greater than 0 denote correct "neural discrimination" between the two contexts and the distance from 0 provides an index of discriminability on that trial. We refer to the trial-by-trial vector of such values to as the multivariate pattern discriminability trial-series (for a related approach to functional connectivity see: Rissman et al., 2004). For each participant, we calculated z[r] Spearman's rank correlation between the multivariate pattern discriminability trial-series in the hippocampus, PHC, and RSC/PCC. We averaged the resultant values to obtain the empirical group mean z[r] Spearman's rank correlations. To assess significance, we used a permutation approach in which we randomized the order of one of the ROI's multivariate pattern discriminability trial-series within each run. We used within-run permutations, rather than permuting the entire trial-series, to mitigate the possibility that between-run differences would artificially reduce the permuted correlations. We then calculated z[r] Spearman's rank correlation between the randomized ROI's vector and the intact ROI's vector. This procedure was carried out 10,000 times per subject. Group-level permutation analysis was conducted by averaging the permuted distributions across subjects, and p values were obtained using Equation 1. In

83

this application, the permutations were independent across subjects because the trial order was independent across subjects.

3.2.9 Multidimensional scaling analysis

To further investigate whether RSC/PCC carries information about items in context in Experiment 1, we performed multidimensional scaling, which is a data-driven, data-reduction approach that allows visualization of the major components of similarity matrices (Kriegeskorte et al., 2008a,b). We first generated a symmetrical representational similarity matrix by averaging values across the diagonal of the matrix (i.e., the same pairs of events across splits). We then converted the matrix to correlation distance (1-z[r]). We extracted the lower triangle of the correlation distance matrix and performed multidimensional scaling using the criterion of metric stress (Kriegeskorte et al., 2008a,b) using package cmdscale in R. We used custom-written code to place the center of the object pairs at the coordinates calculated by the multidimensional scaling procedure. The pictures of the videos were placed on opposite sides of Dimension 1 for visual purposes—i.e., the locations of the videos are arbitrary.

3.2.10 Relationship between representations and behavioral performance on the associative memory task

In Experiment 1, we investigated the relationship between representations and behavioral performance. For each participant, we defined model fit as the z[r] Spearman's rank correlation between their similarity matrix and our proposed model. We calculated the proportion of correct responses during functional scanning within each participant, excluding trials in which the participant did not respond within the response window. We calculated Spearman's rank correlation coefficient between model fit and proportion correct. We used

Spearman's rank correlation because it does not require the assumption that the two variables are normally distributed (as opposed to Pearson's correlation). To assess significance, we calculated the t statistic using the following equation (Krzanowski, 2000):

$$t = r\sqrt{\frac{k-1}{1-r^2}} \tag{3.3}$$

where r is Spearman's rank correlation, and k is equal to $n-1$. We obtained a p value from Student's t distribution with $k-1$ degrees of freedom (20-1-1=18).

To mitigate the possibility of a spurious effect of motion (Power et al., 2012) on the observed relationship between behavioral performance and model fit, we performed a follow-up analysis using a partial correlation approach. Specifically, we examined the relationship between model fit and behavioral performance while holding the effect of head motion constant. Previous reports have used mean motion—defined as the mean amount of motion between successive frames based on the sum of the 3 translation parameters—as a measure of head motion (Van Dijk et al., 2012). Mean motion has previously been shown to be strongly correlated with the total number of motion events and it has been shown to be a reliable index of subject-specific head motion (Van Dijk et al., 2012). We used a related measure, mean framewise displacement, as our measure of head motion. As mentioned above, framewise displacement is the sum of motion across all 6 alignment parameters (3 translation and 3 rotation parameters) between successive frames (Power et al., 2012). To calculate the partial Spearman's rank correlation between the model fit and the proportion correct while holding the effect of head motion constant, we used the following equation (Krzanowski, 2000):

$$r_{xy.z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1-r_{xz}^2)(1-r_{yz}^2)}} \tag{3.4}$$

where x represents the model fit array, y represents the proportion correct array, and z represents the mean framewise displacement array. In our application, r represents Spearman's

rank correlation. The t statistic was calculated using Equation 3. With one variable held constant (i.e., z), $k = n - 1 - 1$. Therefore, the statistical test for partial correlation is the same as in typical correlation analysis with fewer degrees of freedom (Krzanowski, 2000).

3.3 Results

3.3.1 Experiment 1

Investigation of the representation of context

To test the hypothesis that the hippocampus, RSC/PCC, and PHC are involved in context representation, we generated a context matrix, which contains uniformly large values for events that shared the same context and uniformly small values for events that contain different contexts (Fig. 3.1B). We calculated Spearman's rank correlation coefficient between each participant's similarity matrix and the context matrix, and we Fisher's r-to-z transformed (z[r]) the resultant value. We used group-level two-tailed nonparametric p values

Figure 3.3: Investigation of representations in the hippocampus, PHC, and RSC/PCC. A) Average correlation matrices. B) Permutation analysis revealed a significant relationship between the context matrix and the PHC and RSC/PCC similarity matrices (p's < 0.0001) but the relationship failed to reach significance for the hippocampus similarity matrix ($p = 0.71$). C) A searchlight analysis within the hippocampus revealed a significant cluster in the left posterior hippocampus. D) Informational correlativity analysis revealed a significant relationship between all three regions (all p's < 0.0001; sl-HIPP = hippocampus searchlight analysis cluster). E) Permutation analysis revealed a significant relationship between the RSC/PCC similarity matrix and the item-in-context matrix. F) Multidimensional scaling (MDS) analysis provided further evidence for item-in-context representations in RSC/PCC. The locations of the video images are arbitrary other than the left/right assignment while the center of the object pairs are placed at the coordinates determined by the MDS procedure. G) Proposed model of RSC/PCC representations. H) There was a significant relationship between model fit in RSC/PCC and performance on the task (Spearman's rank correlation = 0.51, $t_{18} = 2.51$, $p < 0.05$).

(Ernst, 2004) to assess significance (for related one-tailed approaches to classification analysis see: Chen et al., 2011; Liang et al., 2013b; Stelzer et al., 2013; Etzel, 2015). The PHC and RSC/PCC similarity matrices were significantly related to the context matrix (PHC: $M = 0.15$; RSC/PCC: $M = 0.81$; both $p < 0.0001$; Fig. 3.3A and 3B) while the relationship failed to reach significance for the hippocampus similarity matrix ($M = 0.012$, $p = 0.71$).
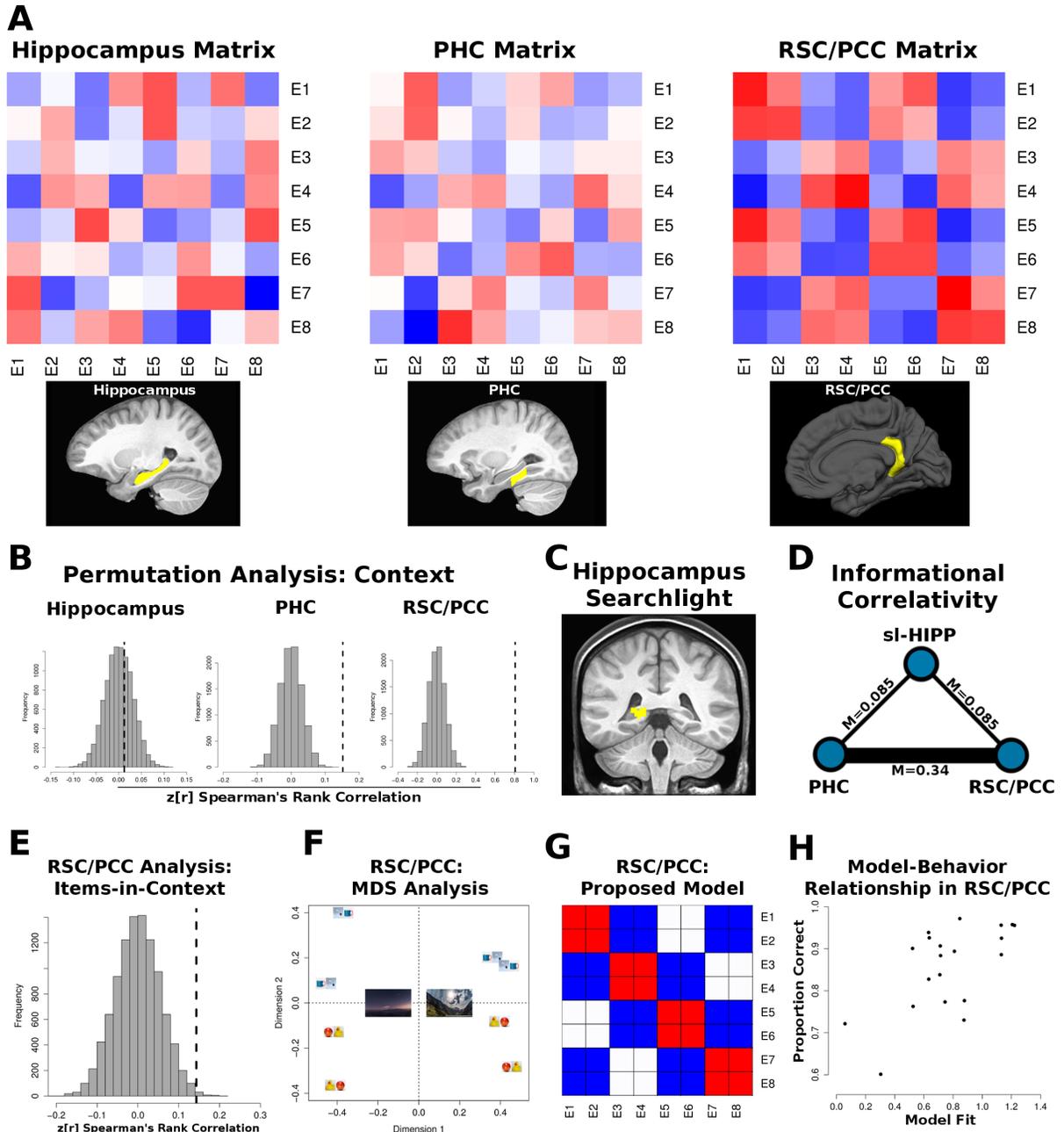
We also tested the hypothesis that PRC carries object information; however, the relationship between the PRC similarity matrix and both the object matrix and the context matrix failed to reach significance (object: $M = 0.032$, $p = 0.30$; context: $M = 0.032$, $p = 0.32$).

We next performed a searchlight analysis within the hippocampus to investigate whether there was a relationship to the context matrix in a portion of the hippocampus, which may have been obscured by uninformative voxels. We used a searchlight radius of 3 voxels within the *sphere_searchlight* function in PyMVPA. A significant cluster was observed in the left posterior hippocampus (33 voxel cluster, parametric voxel-wise threshold $p < 0.05$; Fig. 3.3C). A follow-up analysis, in which we warped the searchlight cluster mask to each participant's native space and ran an ROI analysis similar to the main analysis, indicated that the cluster itself was significantly related to the context matrix ($p = 0.017$; see Etzel et al., 2013). As discussed by Etzel et al. (2013), the follow-up analysis is circular, however it is required in order to conclude that the cluster itself is informative; therefore, the significant effects from this analysis bolster the claim that the hippocampal cluster is significantly related to the context matrix.

To eliminate the possibility that the results in PHC and RSC/PCC were driven solely by strong relationships between an event and "itself"—i.e., because the identity matrix is weakly correlated to the context matrix—we performed a control analysis in which we excluded the entries from the main diagonal of the matrix. The results maintained in both PHC and RSC/PCC (PHC: $M = 0.15$; RSC/PCC: $M = 0.78$; both $p < 0.0001$) and in the hippocampal searchlight cluster ($p < 0.01$). Conversely, the similarity matrix in a control region—the left precentral gyrus (primary motor cortex)—failed to exhibit any sign of a relationship ($M = 0.0072$; $p = 0.85$); instead, it was significantly related to the correct response matrix ($M = 0.26$, $p < 0.0001$; Fig. 3.4). In contrast, the relationship between the PHC and RSC/PCC similarity matrices and the correct response matrix failed to reach significance (PHC: $M = -0.0055$, $p = 0.86$; RSC/PCC: $M = 0.029$, $p = 0.73$; Fig. 3.4). These results
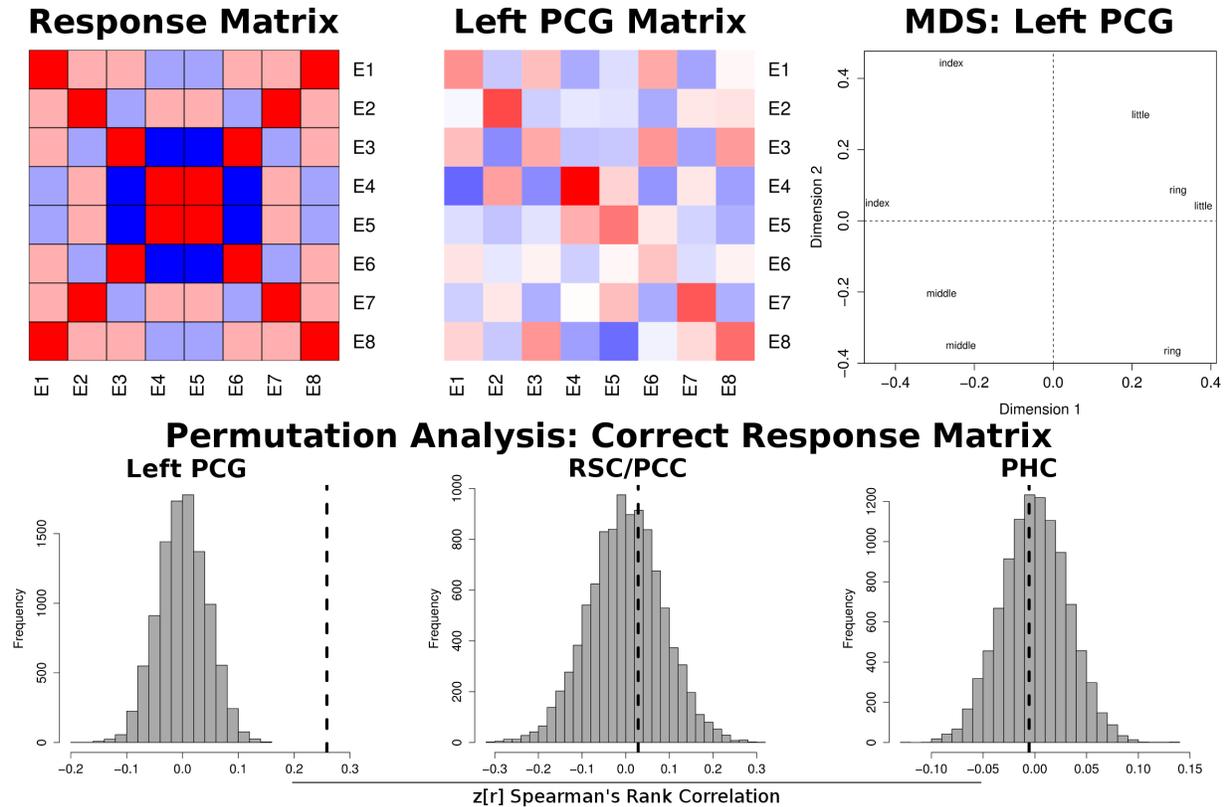
Figure 3.4: Investigation of the relationship to the correct response matrix. Permutation analysis revealed a significant relationship between the left precentral gyrus (lPCG) similarity matrix and the correct response matrix ($M = 0.26$, $p < 0.0001$). There was no sign of a relationship between the correct response matrix and the RSC/PCC and PHC similarity matrices (RSC/PCC: $M = 0.029$, $p = 0.73$; PHC: $M = -0.0055$, $p = 0.86$).

eliminate the possibility that: 1) the relationship between the similarity matrices and the context matrix was driven by correlations between patterns of activity in response to an event and "itself", 2) all cortical regions contain information about context, 3) the relationships to the context matrix were confounded by key-press differences between events in opposing contexts.

We next tested the hypothesis that PHC, RSC/PCC, and the hippocampal searchlight cluster (sl-HIPP) contain similar representations of context on a trial-by-trial basis and thus are related in their processing of contextual information. To address this question, we performed a variant of "informational connectivity" (Coutanche and Thompson-Schill, 2013; Huffman

and Stark, 2014), which we refer to here as informational correlativity. Briefly, we first extracted individual trial estimates of activity (Turner et al., 2012). We then calculated how much like both contexts the pattern of activity was on each trial. We then created a trial-by-trial time-series (for related approaches see: Rissman et al., 2004; Huffman and Stark, 2014), in which the value for each trial was set to the correlation to the same context minus the correlation to the other context (Coutanche and Thompson-Schill, 2013). There was a significant relationship, as measured by Spearman's rank correlation, between trial-by-trial context representation in PHC and RSC/PCC ($M = 0.34$, $p < 0.0001$) and both cortical regions and the hippocampus (both $M = 0.085$, $p < 0.0001$; Fig. 3.3D).

Investigation of the representation of items in context

To test the hypothesis that PHC and RSC/PCC contain item-in-context information, we calculated z[r] Spearman's rank correlation between each participant's similarity matrix and the item-in-context matrix. The RSC/PCC similarity matrix was significantly related to the item-in-context matrix ($M = 0.14$, $p < 0.01$; Fig. 3.3E) but the relationship failed to reach significance for the PHC similarity matrix ($M = 0.012$, $p = 0.76$). To investigate whether the RSC/PCC similarity matrix carried information about items irrespective of context, we generated an item-out-of-context matrix which contains uniformly large values for events which shared the same items but different contexts and uniformly small values for events which contain different items and different contexts. The relationship between the RSC/PCC similarity matrix and the item-out-of-context matrix failed to reach significance ($M = 0.047$, $p = 0.32$), suggesting that RSC/PCC contains conjunctive representations of items in context. To further investigate whether RSC/PCC contains item-in-context information, we performed multidimensional scaling. Multidimensional scaling reduces high-dimensional data to a few dimensions, providing a straightforward method to visualize relationships within similarity matrices (Kriegeskorte et al., 2008a,b). The first two dimensions captured by mul-

tidimensional scaling were: 1) context, 2) items (Fig. 3.3F). Taken together, these results suggest that RSC/PCC contains conjunctive representations of items bound to the context in which they occur.

Investigation of the relationship between RSC/PCC representations and behavior on the associative memory task

Our results (Fig. 3.3B,E-F) revealed a clear relationship (i.e., $p < 0.01$) between the similarity matrix in RSC/PCC and both the context matrix and the item-in-context matrix. Therefore, we generated a proposed model of representations in RSC/PCC (Fig. 3.3G), which is ranked in descending order of similarity: 1) events that shared the same context and items, 2) events with the same context but different items, 3) events with different contexts. There was a significant relationship between model fit (z[r] Spearman's rank correlation between the RSC/PCC similarity matrix and our proposed model) and proportion correct (Spearman's rank correlation $= 0.51$, $t_{18} = 2.51$, $p < 0.05$; Fig. 3.3H), which maintained when controlling for the effect of head motion (partial Spearman's rank correlation $= 0.47$, $t_{17} = 2.18$, $p < 0.05$), suggesting that the relationship between the model fit of the RSC/PCC similarity matrix and behavioral performance was not influenced by head motion.

Investigation of representations in parahippocampal place area and retrosplenial complex

There has been extensive evidence of scene and context processing in the parahippocampal place area (PPA) and the retrosplenial complex (RS-Complex; e.g., Epstein and Kanwisher, 1998; Epstein et al., 2007; Julian et al., 2012; Vass and Epstein, 2013, 2016; Marchette et al., 2014, 2015). These regions are in close anatomical proximity to the PHC and RSC/PCC ROIs, however the ROIs are largely non-overlapping. Specifically, PPA tends to be located
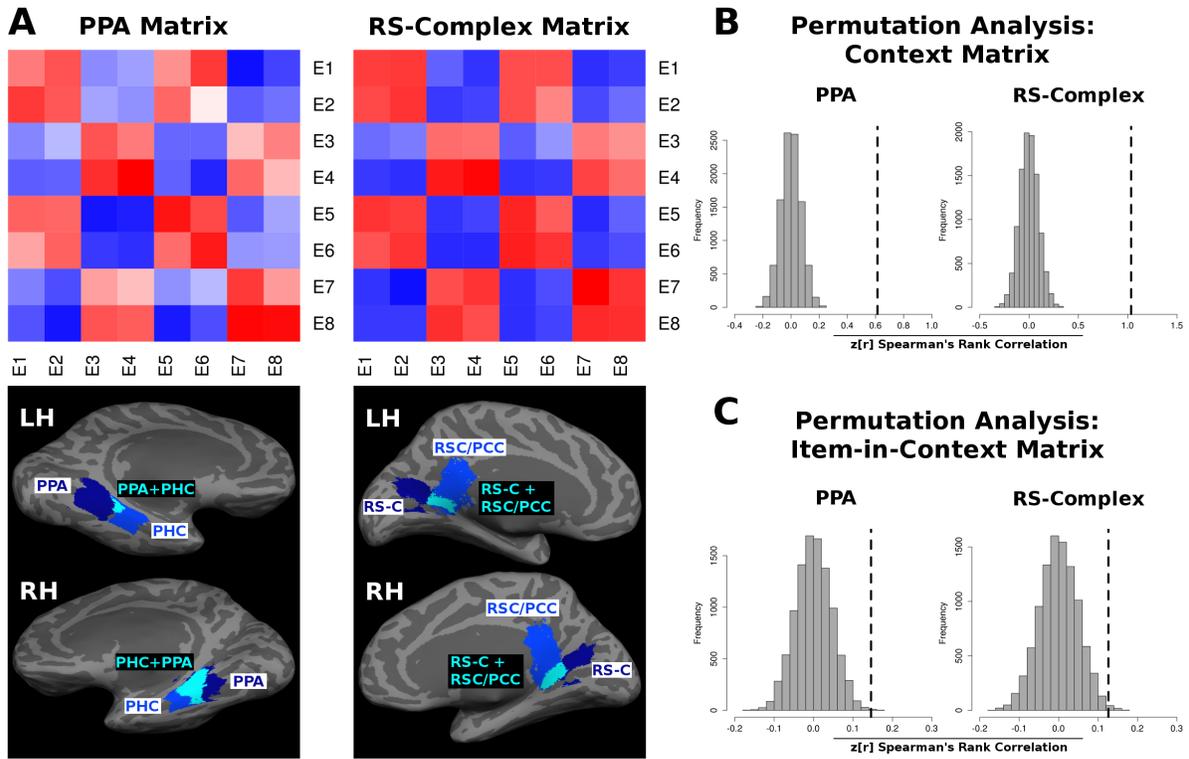
Figure 3.5: Investigation of representations in parahippocampal place area (PPA) and retrosplenial complex (RS-Complex, also called "RS-C" in mask images). A) Average correlation matrices and ROI masks relative to PHC and RSC/PCC (LH = left hemisphere; RH = right hemisphere). B) Permutation analysis revealed a significant relationship between the context matrix and the PPA and RS-Complex similarity matrices (PPA: $M = 0.61$, $p < 0.0001$; RS-Complex: $M = 1.03$, $p < 0.0001$). C) Permutation analysis revealed a significant relationship between the item-in-context matrix and the PPA and RS-Complex similarity matrices (PPA: $M = 0.15$, $p < 0.005$; RS-Complex: $M = 0.13$, $p < 0.01$).

posterior to PHC (i.e., along the parahippocampal gyrus, but posterior to the landmarks like the splenium of the corpus callosum) and retrosplenial complex tends to be located posterior to RSC/PCC (anatomical boundaries are shown in Fig. 3.5A). Within the anatomical masks for PPA and RS-Complex, we selected the 100 most active voxels from each hemisphere using a contrast of events greater than perceptual baseline. We combined the masks into bilateral PPA and RS-Complex. Similar to the results in RSC/PCC, the PPA and RS-Complex similarity matrices were significantly related to the context matrix (PPA: $M = 0.61$, $p < 0.0001$; RS-Complex: $M = 1.04$, $p < 0.0001$; Fig. 3.5B) and the item-in-context matrix (PPA: $M = 0.15$, $p < 0.005$; RS-Complex: $M = 0.13$, $p < 0.01$; Fig. 3.5C).
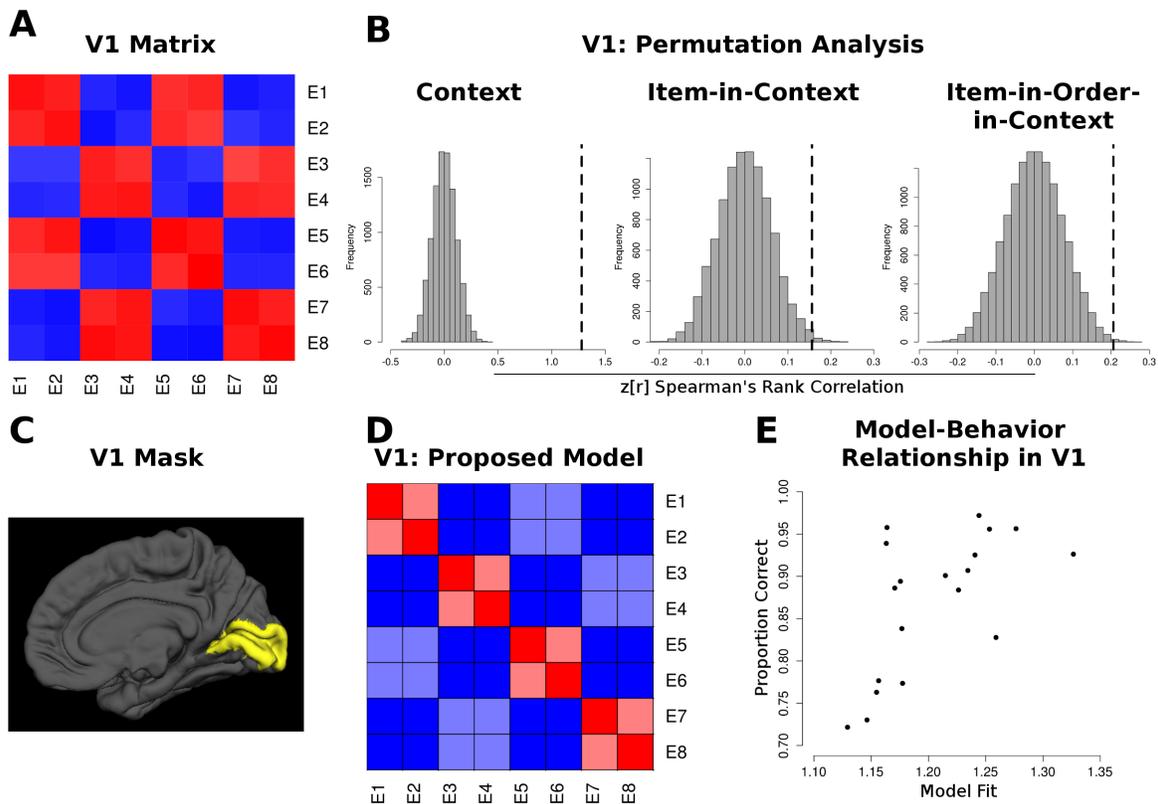
Figure 3.6: Investigation of representations in V1. A) Average correlation matrix. B) Permutation analysis revealed a significant relationship between the V1 similarity matrix and the context matrix ($M = 1.23$, $p < 0.0001$), the item-in-context matrix ($M = 0.16$, $p < 0.01$), and the item-in-order-in-context matrix ($M = 0.21$, $p < 0.005$). C) V1 mask. D) Proposed model of V1 representations. E) There was a significant relationship between model fit and performance on the task (Spearman's rank correlation = 0.58, $t_{17} = 2.97$, $p < 0.01$).

Investigation of representation in a control visual region: V1

Given that we used visual stimuli, we aimed to determine whether the above findings were limited to classically-defined scene and context regions or whether we would find similar effects in early visual areas. We ran the same analysis using V1 as an ROI. We found a similar pattern of results to RSC/PCC, PPA, and RS-Complex, including a relationship to both the context matrix ($M = 1.23$, $p < 0.0001$) and the item-in-context matrix ($M = 0.16$, $p < 0.01$). Similar to the findings in RSC/PCC, the relationship to the item-out-of-context

matrix failed to reach significance ($M = 0.014$, $p = 0.81$). Notably, we also found evidence for a relationship to the item-in-order-in-context model ($M = 0.21$, $p < 0.005$; Fig. 3.6A-C). We generated a proposed model of representations in V1 (Fig. 3.6D), which is ranked in descending order of similarity: 1) identical events, 2) events with the same stimuli but switched order, 3) events with the same context but different items, 4) events with different contexts. There was a significant relationship between model fit in V1 and proportion correct (Spearman's rank correlation $= 0.58$, $t_{17} = 2.97$, $p < 0.01$; Fig. 3.6E), which maintained when controlling for the effect of head motion (partial Spearman's rank correlation $= 0.58$, $t_{16} = 2.88$, $p = 0.01$; note, one participant was dropped from the analysis due to low model fit, however, the effect was stronger, for both approaches, when the participant was included in the analysis).

Whole-brain searchlight analysis

Finally, we report the results of an exploratory whole-brain searchlight analysis (3 voxel radius) which was conducted to investigate the prevalence of the relationship to each of our model matrices. We ran the searchlight in native space and warped the results to our group template using ANTs. For each contrast, we used a voxel-wise threshold of $p < 0.01$ (parametric) and a cluster threshold of $p < 0.05$ (cluster threshold was determined using Monte Carlo simulation with a simulated blur of 3.75 mm FWHM, i.e., half of the searchlight radius). First, we investigated the relationship to the context matrix. Next, we masked the whole-brain searchlight results for the item-in-context matrix to only include regions that were significantly related to the context matrix. Finally, we masked the whole-brain searchlight results for the item-in-order-in-context matrix to only include regions that were significantly related to both the context matrix and the item-in-context matrix. The overlap map was warped to an inflated brain for visualization using FreeSurfer. The analysis revealed widespread relationships to the context matrix (Fig. 3.7). Additionally, several regions were
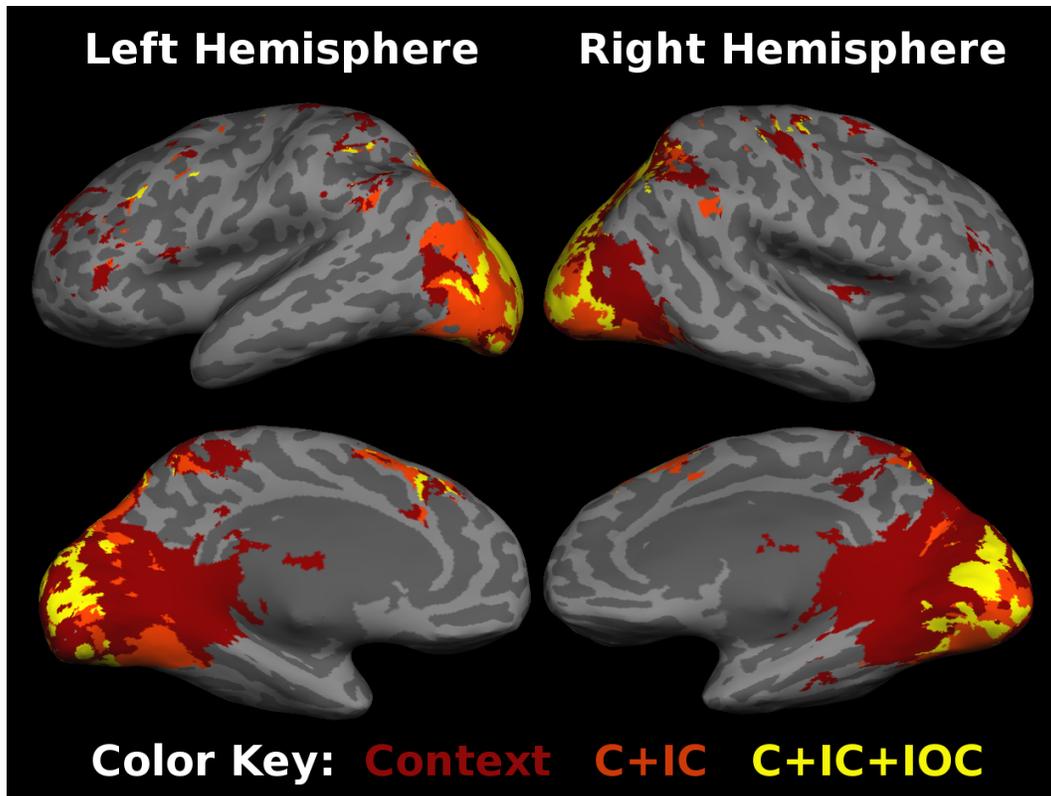
Figure 3.7: Whole-brain searchlight analysis revealed context clusters in the occipital lobe (nearly the entire lobe bilaterally), the temporal lobe, the frontal lobe, and the parietal lobe (shown in red). Additionally, there were distributed clusters that survived the union of context and item-in-context (shown in orange) as well as the union of context, item-in-context, and item-in-order-in-context (shown in yellow).

related to both the context matrix and the item-in-context matrix. Finally, several regions were related to all three matrices. While there were significant clusters in all four lobes, the most prominent findings were in the occipital lobes, where many voxels were significantly related to all three matrices.

Interim Summary

The results from Experiment 1 provide clear evidence for a relationship between the context matrix and the similarity matrices in the left posterior hippocampus, parahippocampal cortex, RSC/PCC, PPA, and RS-Complex—i.e., patterns of activity in these regions could

be used to discriminate between events that contained different videos (our operationally defined contexts). However, our results revealed that patterns of activity in V1 could be used to successfully discriminate between all aspects of events—i.e., the V1 similarity matrix was significantly related to the context matrix, the item-in-context matrix, and the item-in-order-in-context matrix. Moreover, a whole brain searchlight analysis revealed relationships to all three matrices throughout the brain, but the most prominent results were located in the occipital lobes. Taken together, the interpretation of our ROI-based results could be confounded by the low-level visual differences between our stimuli. Our primary aim in Experiment 2 was to investigate whether our ROIs would still carry context information after eliminating low-level differences between our contexts.

3.3.2 Experiment 2

There are many possible approaches to reduce the influence of low-level features, but we chose to investigate whether our ROIs exhibit invariant representations of context (see Discussion for other possible approaches). Specifically, we used multiple images to define our contexts and objects. We also used a combination of stimulus filtering and computational model testing to attempt to eliminate low-level differences between our contexts and objects. Given that the nature of the associative memory task was unchanged, we hypothesized that regions that carry information about contexts and objects—as those terms relate to performance on the associative memory task—should do so in an invariant manner. Conversely, if our ROI-based results from Experiment 1 were influenced by low-level stimulus features, then we should fail to observe a relationship to the context matrix in our ROIs.
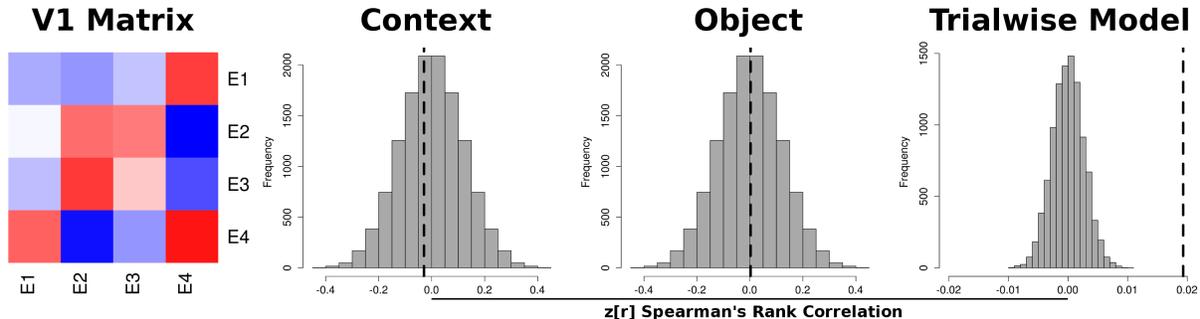
Figure 3.8: The low-level confound was attenuated in Experiment 2. The relationship between the V1 similarity matrix and both the context matrix and the object matrix failed to reach significance (context matrix: $M = -0.029$, $p = 0.81$; object matrix: $M = 0.0025$, $p = 0.99$). Trialwise analysis revealed a significant relationship between the empirical V1 similarity matrix and the model V1 similarity matrix ($M = 0.019$, $p < 0.0001$).

Investigation of representation in V1

First, we investigated whether our stimulus filtering and computational modeling approaches removed the low-level visual confound that was observed in Experiment 1. Importantly, the V1 similarity matrix showed no sign of a relationship to either the context matrix ($M = -0.029$, $p = 0.81$) or the object matrix ($M = 0.0025$, $p = 0.99$; Fig. 3.8). Next, a positive control analysis was performed in order to demonstrate a relationship between the empirical V1 similarity matrix and the model V1 similarity matrix. We conducted a trialwise analysis which enabled us to model activity in response to both the scene image and the object image on each trial. Specifically, the V1 model was "shown" the same stimulus sequence as the participant. On each trial, the V1 model response vectors were extracted in repsonse to the scene image and in response to the scene and object images. The resultant vectors were combined using weighted averaging (4/5 of the scene only vector plus 1/5 of the scene plus object vector). The mean pattern of activity was removed within the odd and even splits, and correlation matrices were generated by correlating the pattern of activity across all of trials across the odd and even splits. A similar approach was conducted to generate the empirical similarity matrix (also see: Informational correlativity analysis). We calculated Spearman's rank correlation between the trialwise empirical V1 similarity matrix and the model V1
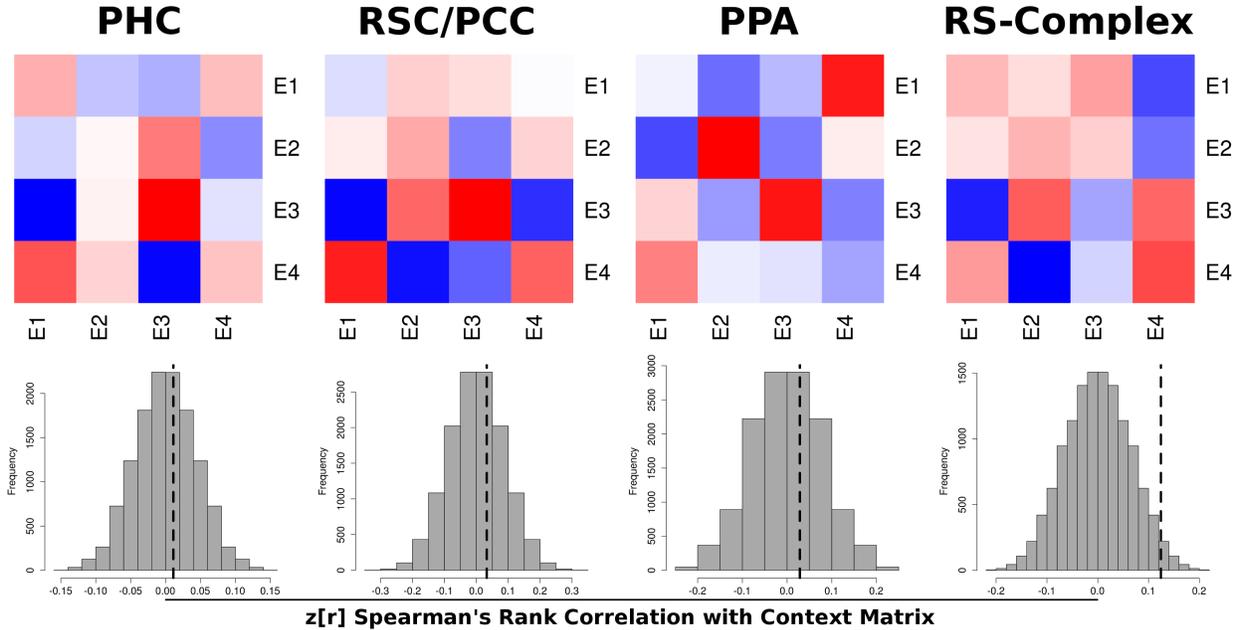
Figure 3.9: Investigation of invariant context representation. The relationship between the context matrix and the PHC, RSC/PCC, PPA, and RS-Complex similarity matrices failed to reach significance (PHC: $M = 0.011$, $p = 0.80$; RSC/PCC: $M = 0.033$, $p = 0.71$; PPA: $M = 0.029$, $p = 0.69$; RS-Complex: $M = 0.12$, $p = 0.051$).

similarity matrix using the motion censoring steps described for representational correlativity analysis. Trialwise analysis revealed a significant relationship between the empirical V1 similarity matrix and the model V1 similarity matrix ($M = 0.019$, $p < 0.0001$; Fig. 3.8). Note, the relationship between the two matrices was numerically small but highly reliable, which was expected given that individual trial estimates of activity were used for the analysis. We also observed a significant relationship between the V1 similarity matrix and the response matrix ($M = 0.38$, $p < 0.005$), which was likely driven by the hemifield differences of the selected responses.

Investigation of the representation of context

The results in V1 suggest that the low-level visual confound has been, at the very least, attenuated in Experiment 2. We next investigated whether the relationship to the context

98

matrix maintained in our scene and context processing ROIs. Permutation analysis revealed a severely diminished relationship between the PHC, RSC/PCC, PPA, and RS-Complex similarity matrices and the context matrix relative to Experiment 1 (PHC: $M = 0.011$, $p = 0.80$; RSC/PCC: $M = 0.033$, $p = 0.71$; PPA: $M = 0.029$, $p = 0.69$; RS-Complex: $M = 0.12$, $p = 0.051$; Fig. 3.9). Thus, reducing (or eliminating) the low-level perceptual differences across contexts markedly decreased the substantial contextual effects that we observed in Experiment 1.

Activation analysis: Events versus perceptual baseline

To ensure that the data were reliable and that the task was significantly activating our ROIs, a standard activation analysis (events vs perceptual baseline) was conducted. This revealed significantly greater BOLD activity for events than the perceptual baseline task in PHC, RSC/PCC, PPA, and RS-Complex (PHC: $t_9 = 6.4443$, $p < 0.0005$; RSC/PCC: $t_9 = 2.9574$, $p < 0.02$; Left anatomical PPA: $t_9 = 7.3008$, $p < 0.0001$; Right anatomical PPA: $t_9 = 6.6446$, $p < 0.0001$; Left anatomical retrosplenial complex: $t_9 = 4.5301$, $p < 0.005$; Right anatomical retrosplenial complex: $t_9 = 4.2866$, $p < 0.005$). Additionally, a whole-brain analysis revealed extensive activation throughout the ventral temporal lobe (data not shown). These findings suggest that all of our scene ROIs responded to the events relative to the baseline task even though there was little evidence for a relationship to the context matrix.

Investigation of object representation

To test the hypothesis that PRC carries object-level information (even when the low-level stimulus features have been matched), we generated an object matrix, which contains uniformly large values for events that share the same objects and uniformly small values for events that contain different objects (Fig. 3.2B). Permutation analysis revealed a significant

relationship between the PRC similarity matrix and the object matrix ($M = 0.28$, $p = 0.002$), while the relationship to the context matrix failed to reach significance ($M = -0.022$, $p = 0.81$; Fig. 3.10A). We next tested the prevalence of the relationship to the object matrix using a whole-brain searchlight analysis (3 voxel radius). A significant cluster was observed in left PRC, supporting the ROI-based results (57 voxel cluster, parametric voxel-wise $p < 0.01$; Fig. 3.10B). Additionally, two clusters in the right anterior temporal cortex were observed (53 and 42 voxel clusters; middle and left panels of Fig. 3.10B, respectively). A follow-up analysis revealed that all three clusters themselves were significantly related to the object matrix (cluster 1: $M = 0.44$, $p < 0.0005$; cluster 2: $M = 0.38$, $p < 0.0005$; cluster 3: $M = 0.40$, $p = 0.011$, one participant was excluded from the cluster 3 analysis due to insufficient coverage). As discussed by Etzel et al. (2013), the follow-up analysis is circular, however it is required in order to conclude that the cluster itself is informative; therefore, the significant effects from this analysis bolster the claim that the clusters are significantly related to the object matrix.

Testing for a double-dissociation between V1 and PRC

We next performed trialwise analysis to investigate whether there was a relationship between PRC and the V1 model. In contrast to V1, the relationship between the PRC similarity matrix and the model V1 similarity matrix failed to reach significance ($M = 0.0025$, $p = 0.36$; left panel Fig. 3.10C). A nonparametric difference test revealed that the relationship between the empirical V1 similarity matrix and the model V1 similarity matrix was significantly stronger than the relationship between PRC similarity matrix and the model V1 similarity matrix ($M = 0.017$, $p < 0.0001$; middle panel Fig. 3.10C). Finally, we observed some evidence that the relationship between the PRC similarity matrix and the object matrix was stronger than the relationship between the V1 similarity matrix and the object matrix ($M = 0.28$, one-tailed $p < 0.05$; right panel Fig. 3.10C). We would like to emphasize that we advocate for
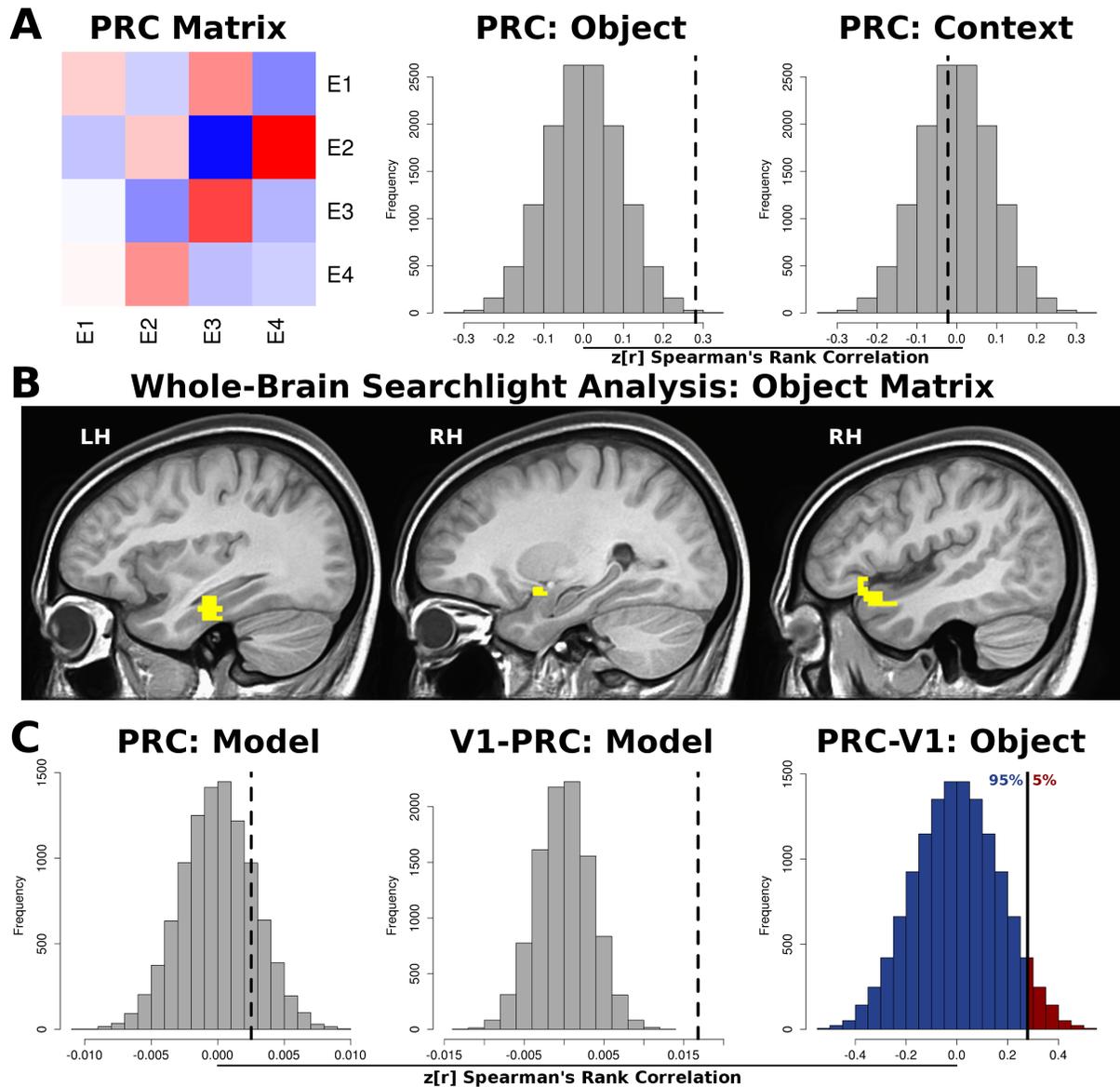
Figure 3.10: Investigation of invariant object representation. A) Permutation analysis revealed a significant relationship between the PRC similarity matrix and the object matrix ($M = 0.28$, $p = 0.002$) but the relationship to the context matrix failed to reach significance ($M = -0.022$, $p = 0.81$). B) A whole-brain searchlight analysis revealed a cluster in the left PRC as well as two other clusters in the right anterior temporal lobe. C) The relationship between the trialwise PRC similarity matrix and the trialwise V1 model similarity matrix failed to reach significance ($M = 0.0025$, $p = 0.36$; left panel), the relationship between the trialwise V1 similarity matrix and the trialwise V1 model similarity matrix was significantly stronger that the relationship between the trialwise PRC similarity matrix and the trialwise V1 model similarity matrix ($M = 0.017$, $p < 0.0001$; middle panel), and there was evidence that the relationship between the PRC similarity matrix and the object matrix was stronger than the relationship between the V1 similarity matrix and the object matrix ($M = 0.28$, one-tailed $p < 0.05$; right panel).

101

the use of two-tailed permutation tests (e.g., Equations 1 and 2); however, to demonstrate feasibility, we report the results of a one-tailed difference test.

3.4   Discussion

Our experiments aimed to investigate whether subregions of the MTL carry information about context, items, order, and their conjunctions. We built upon a well-designed approach used in the rodent (Rajji et al., 2006; Komorowski et al., 2009, 2013; Navawongse and Eichenbaum, 2013; Tort et al., 2013; McKenzie et al., 2014; Farovik et al., 2015; Keene et al., 2016) and observed many analogous patterns using fMRI in humans. Specifically, we observed clear evidence of a relationship between the context matrix and the similarity matrices in the posterior hippocampus, PHC, RSC/PCC, PPA, and RS-Complex, however such information was also present in V1. Controlling for low-level visual differences across contexts markedly diminished the relationships to the context matrix. These findings raise important questions concerning not only how we attempt to understand or define *context*, but also how we attempt to decode information from brain regions to infer what they represent. While our experimental design in Experiment 2 tested for an abstract notion of context—i.e., invariant representations across different instances of the context—it need not be the case that context is defined in such an abstract manner. In fact, there are many other possibilities for experimental designs that aim to investigate context and object representation. We will expand upon these issues below.

3.4.1 Investigation of the representation of distinct contexts and distinct objects

In Experiment 1, we aimed to create a human version of the context-guided object association task to investigate representations in the hippocampus, PHC, and RSC/PCC. We showed that both the PHC and the RSC/PCC similarity matrices were significantly related to the context matrix. The results of a within-hippocampus searchlight analysis revealed a cluster of voxels that were significantly related to the context matrix in the left posterior hippocampus, and a follow-up analysis revealed that the cluster itself was significantly related to the context matrix. Additionally, by showing that the hippocampus cluster, PHC, and RSC/PCC contain similar representations of context on a trial-by-trial basis, our results support the hypothesis that the hippocampus, PHC, and RSC/PCC are intimately related in processing contextual information (Vann et al., 2009; Aggleton, 2010; Ranganath and Ritchey, 2012; Bucci and Robinson, 2014). These results extend our previous finding of stimulus-dependent informational correlativity between RSC/PCC and PHC from the representation of different categories (e.g., faces and scenes; Huffman and Stark, 2014) to the representation of individual contexts. There are many reasons to expect that fMRI data will be noisier than neurophysiological data in the hippocampus. For example, previous reports have suggested a lack of topographic organization of place cells in the hippocampus (Redish et al., 2001) and each of our voxels contains thousands of cells. The largest differences in patterns of activity in the rodent dorsal hippocampus (the homolog of human posterior hippocampus; McKenzie et al., 2014) and medial entorhinal cortex (Keene et al., 2016) were observed in response to changes to the context; therefore, the observed relationship to the context matrix in the posterior hippocampus is consistent with the previous reports. Moreover, our findings extend the results from the rodent hippocampus and medial entorhinal cortex to the human hippocampus, PHC, and RSC/PCC.

We next showed that that in addition to its relationship to the context matrix, the RSC/PCC

similarity matrix was related to the item-in-context matrix. Multidimensional scaling suggested that the first two dimensions represented within the RSC/PCC similarity matrix were the event context followed by the event items. These results suggest that RSC/PCC contains conjunctive item-in-context representations when the item and context stimuli are distinct. These findings support studies in the rodent that demonstrated that RSC is necessary for performance on tasks that require the formation of associations between items and locations (Ennaceur et al., 1997; Vann and Aggleton, 2002; Parron and Save, 2004). Furthermore, our results extend to human RSC/PCC the finding that the medial entorhinal cortex carries conjunctive item-in-context information in addition to spatial information (e.g., context) as rats performed the context-guided object association task (Keene et al., 2016).

Our results provide clear evidence for a relationship between the RSC/PCC similarity matrix and both the context matrix and the item-in-context matrix. We next asked whether this relationship is related to task performance. We generated a proposed model in which representations are hypothesized to be maximally similar for events that share the same context and items, second most similar for events that take place in the same context but with different items, and least similar for events that take place in a different context (Fig. 3.3G). There was a significant relationship between RSC/PCC model fit and behavioral performance (as measured by Spearman's rank correlation), and the effect maintained when controlling for the potentially confounding effect of head motion, which is at least consistent with a role for RSC/PCC in associative memory performance (Ennaceur et al., 1997; Vann and Aggleton, 2002; Parron and Save, 2004; Bucci and Robinson, 2014).

We next investigated representations in PPA and RS-Complex, which have been extensively studied in human neuroimaging studies of the processing of scenes and spatial information (e.g., Epstein and Kanwisher, 1998; Epstein et al., 2007; Morgan et al., 2011; Julian et al., 2012; Vass and Epstein, 2013, 2016; Marchette et al., 2014, 2015). As mentioned above, PPA and RS-Complex are in close anatomical proximity to PHC and RSC/PCC, respec-

tively; however, the ROIs are largeley independent (Fig. 3.5A) and PPA and RS-Complex are functional ROIs whereas PHC and RSC/PCC are anatomical ROIs. The results in PPA and RS-Complex were similar to our findings in RSC/PCC. Specifically, the similarity matrices in both regions were significantly related to both the context matrix and the item-in-context matrix. As we will subsequently discuss, a critical question arises about whether our results are dependent on low-level differences between the stimuli; however, whether or not the representations in RSC/PCC, PPA, and RS-Complex are influenced by low-level visual features, our results suggest that these regions carry information not only about scenes or contexts but also, at least under certain conditions, about item-in-context information. Similarly, even if previous results in the rodent (Komorowski et al., 2009, 2013; Navawongse and Eichenbaum, 2013; Tort et al., 2013; McKenzie et al., 2014; Keene et al., 2016) are dependent on low-level differences between the stimuli, it does not detract from the conclusion that the hippocampus and MTL cortical regions have access to all of the necessary components for "what-where" processing. As we will discuss in more detail later, however, such a finding could hinder the interpretation of the organization of representations in the MTL.

The V1 similarity matrix was related to the context matrix and the item-in-context matrix, suggesting that our results in RSC/PCC, PPA, and RS-Complex could have been influenced by low-level visual differences between the stimuli. The V1 similarity matrix was also related to the item-in-order-in-context matrix, suggesting that early processing areas can exhibit distinct patterns of activity in response to a reconfiguration of the same stimuli. Moreover, we observed a significant relationship between representations in V1 and proportion correct. We cannot rule out the possible contribution of attention or other task factors, however we suggest that these results could be interpreted in at least two ways. First, significant brain-behavior relationships are suggestive but not sufficient to conclude that a given brain region is involved in mnemonic processing. Second, the significant brain-behavior correlation is at least consistent with the notion that V1 could play a role in associative memory

(cf. Weinberger, 2004). Future studies which employ pre-training and post-training scanning in addition to including control stimuli (e.g., identically structured events with an equal number of exposures but without a learned behavioral response or outcome) will be necessary to elucidate the possible mnemonic contribution of V1 and RSC/PCC; however, even in these cases it will be difficult to rule out between-condition differences in attention. Relatedly, previous findings in the rodent suggest that the hippocampus develops conjunctive item-in-position neurons over the course of training on the context-guided object association task (Komorowski et al., 2009; Tort et al., 2013). These results are important because they show changes to representations in the hippocampus as a result of learning, however it is not known how learning alters representations in cortical regions nor can it be ruled out that training causes an animal to form more stereotyped behavior or expectations which could alter representations in regions outside of the MTL. Altogether, the results in V1 highlight the difficulty of disentangling the role of a brain region in memory versus processing, especially given that MTL regions, and the hippocampus in particular, receive inputs from all sensory modalities (e.g., in addition to the well-known modulation of hippocampal firing based on visual cues, hippocampal neurons have been shown to be modulated by auditory cues: Itskov et al., 2012; gustatory cues: Ho et al., 2011; olfactory cues: Zhang and Manahan-Vaughan, 2015; and tactile cues: Gener et al., 2013).

We performed an exploratory whole-brain searchlight analysis to investigate the prevalence of the relationship to each of our matrices. Our results revealed relationships to the context matrix throughout the brain. Additionally, there were distributed regions that were related to both the context and the item-in-context matrix as well as to all three matrices. Notably, the strongest effects were observed in the occipital lobes, where many voxels were related to not only the context matrix but also the item-in-context matrix and the item-in-order-in-context matrix. In hindsight this result is not surprising given the visual nature of the events, but it could provide a challenge to the conclusion that context plays an organizing role for processing within the MTL (McKenzie et al., 2014, 2015; Farovik et al., 2015; Keene et al.,

2016). In the rodent context-guided object association task, the contexts differed in terms of visual and tactile cues and the items differed in terms of the digging media and the olfactory cues. Because the stimuli were changed so dramatically and in a relatively unconstrained manner it is difficult to interpret the degree to which regions differed in their responses. In Experiment 1, we used distinct video clips for our context (full screen display) and distinct images of objects for our item stimuli (150 by 150 pixels), thus the physical differences between our contexts were larger than the physical differences between our objects (as was the case in the degree of differences in the rodent experiments). We suggest that one way to demonstrate that context plays an organizing role of representations in the MTL would be to manipulate the contexts and objects to a physically similar degree within the same modality (e.g., visual features). If the results are similar under these conditions, then it would bolster the argument that the MTL organizes events into a context-based schema (McKenzie et al., 2015), largely eliminating the potential concerns raised here.

Previous studies have shown differential place cell firing in the rat hippocampus in response to the same physical stimuli as a result of behavioral conditions (Wood et al., 2000; Smith and Mizumori, 2006b) and expectations (Skaggs and McNaughton, 1998; Allen et al., 2016), suggesting that psychological context can play a dramatic role even within the same physical environment (for reviews see: Smith and Mizumori, 2006a; Smith and Bulkin, 2014). A human neuroimaging study used a sequence task to investigate the influence of temporal context on representations in the hippocampus (Hsieh et al., 2014). They reported significantly more similar patterns of activity in response to the same item in the correct order than in response to the same item in a random order, suggesting that temporal context stabilizes hippocampal representations. A related study in the rodent observed cells in the hippocampus that carried information about temporal context (Allen et al., 2016). Designs that manipulate temporal context are interesting because they allow investigation of the influence of context without manipulating the physical stimulus. These examples suggest that context undoubtedly plays an important role in the organization of representations in the

hippocampus, however future experiments should test the degree to which physical versus psychological factors influence representations in both rats and humans. Notably, studies that have recorded simultaneously from V1 and the hippocampus have shown "place cell" activity in V1 (Ji and Wilson, 2007) which was also shown to precede place cell activity in the hippocampus (Haggerty and Ji, 2015). Thus, a more complete understanding of the sensory versus cognitive influences on representations in the MTL can be realized by studies that incorporate simultaneous recordings in sensory cortices and the MTL. Finally, we argue that developing cross-species memory tasks (as we have attempted here) is vital to furthering our understanding of the neurobiology of learning and memory, but the present results indicate that we need to be cautious when designing visual tasks for humans. For example, human neuroimaging studies have indicated that patterns of activity in early visual cortex can be used to successfully decode landmark identity (Morgan et al., 2011; Marchette et al., 2015), view within a city (i.e., images of same location and same facing direction; Vass and Epstein, 2013), and the visual cues that are present in a virtual environment (Op de Beeck et al., 2013). We suggest that image manipulation and model testing of the stimulus set before running an experiment can provide one method for avoiding a low-level visual confound. We provide an example of this approach in Experiment 2.

3.4.2 Investigation of invariant context representation

In Experiment 2, we aimed to test the extent to which the results from Experiment 1 were influenced by low-level differences between the stimuli. We used a combined approach of image manipulation (histogram matching) and computational model testing to eliminate the presence of category information from the low-level visual features. In stark contrast to the results of Experiment 1, there was no sign of a relationship between the V1 similarity matrix and both the context matrix and the object matrix. These results suggest that we adequately reduced the low-level visual confound in this stimulus set. Importantly, two

control analyses showed that V1 contained task-relevant information. First, a trialwise analysis revealed a relationship between the empirical V1 similarity matrix and the model V1 similarity matrix. Second, there was a significant relationship between the V1 similarity matrix and the response matrix, which likely reflects the hemifield differences based on the left versus right response (see Fig. 3.2A). The fact that there was a relationship between the response matrix and the V1 similarity matrix provides further evidence to suggest that the low-level confound was attenuated—i.e., these results suggest that the image information was canceled out and that the only statistically reliable visual information at the level of between run analysis was the filling in of the box in either hemifield. Taken together, the positive control analyses in V1 establish data quality.

We next investigated whether PHC, RSC/PCC, PPA, and RS-Complex contain invariant representations of context. We found a severely diminished relationship between the similarity matrices in these regions and the context matrix relative to Experiment 1. However, we do not want to strongly interpret these relatively null results. While the effect was clearly diminished, we cannot say it was entirely eliminated or that it would not be observed under other circumstances. For example, it is possible that our sample size was too small to detect significant relationships in these regions. Furthermore, it is possible that representations exist on a more fine-grained level than is afforded in our current fMRI experiment. Future studies using neurophysiology and fMRI will be useful to help elucidate the conditions under which these regions exhibit invariant context coding. For example, Marchette et al. (2015) investigated whether there were stable representations between indoor and outdoor images of landmarks. They reported invariant representations across indoor and outdoor images of the same landmark in both PPA and RS-Complex. Moreover, they provided evidence to suggest that such representations might depend on landmark familiarity in PPA. Importantly, pixelwise correlation, HMAX, GIST, and early visual cortex failed to exhibit invariant representations across indoor and outdoor images, suggesting that low-level visual features were not sufficient for indoor-outdoor generalization. Future studies which manipulate stimulus

familiarity will be useful for understanding the involvement of the MTL, RSC, PPA, and RS-Complex in context representation as well as the role of memory in the formation of these representations. For example, it is possible that our participants did not have real-world experience with Saint Peter's Basilica or with the U.S. Capitol Building which led to little evidence for invariant context representation. Importantly, all of our scene ROIs showed greater BOLD activation for events compared to the perceptual baseline task, suggesting that they carried information about events compared to the baseline.

3.4.3 Investigation of invariant object representation

In contrast to the relative lack of findings for invariant context representation, we observed a significant relationship between the PRC similarity matrix and the object matrix in Experiment 2. These results strongly support the hypothesis that PRC is involved in the representation of objects. A whole-brain searchlight analysis revealed three significant clusters, the first of which was centered in left PRC and the other two of which were located in right anterior temporal lobe (one of which was in close proximity to PRC, containing some overlapping voxels). Importantly, a follow-up analysis revealed that the clusters themselves were significantly related to the object matrix. These results corroborate the ROI-based approach and suggest that the relationship to the object matrix is relatively exclusive to the more anterior portions of the MTL.

We propose that an important next step is to confirm that there is a dissociation between representations in early visual areas (e.g., V1) and the representations in PRC, which mitigates the possibility that representations were inherited from earlier processing regions. Accordingly, we tested for a double dissociation between representations in V1 and in PRC. In our analysis, we used nonparametric difference tests to investigate representational differences between there regions. Specifically, after establishing that PRC was not related to the

trialwise model V1 similarity matrix, we showed that the relationship between the empirical V1 similarity matrix and the model V1 similarity matrix was stronger than the relationship between the PRC similarity matrix and the model V1 similarity matrix. Next, we provided evidence to suggest that the relationship between the PRC similarity matrix and the object matrix was stronger than the relationship between the V1 similarity matrix and the object matrix. These results provide evidence for higher-level representations in PRC than in V1 and suggest that PRC contains invariant car key and house key representations across multiple exemplars. Future studies could investigate whether these effects are dependent on the associative memory task.

An alternative to the difference tests that we used in the present report would be to use a partial correlation approach. For example, Clarke and Tyler (2014) used a partial correlation approach to investigate representations throughout the visual stream. One benefit of the partial correlation approach is that it is straightforward to implement in experiments with larger and less constrained stimulus sets than we used in Experiment 2. In fact, Clarke and Tyler (2014) used a stimulus set of 131 unique objects to investigate object representation. The results of their partial correlation representational similarity analysis suggest that PRC represents fine-grained semantic information about individual objects. Importantly, this effect maintained while holding the effect of a model of V1 representations constant (among other models). Futhermore, they used a modeling approach to show that BOLD activation in PRC was modulated by the confusability of objects. Taken together, their results suggest a role for PRC in fine-grained semantic representations of objects (for review see: Clarke and Tyler, 2015). Similarly, research in patient populations has revealed a necessary role for the anterior temporal cortex, and PRC in particular, in naming highly confusable objects (Kivisaari et al., 2012; Wright et al., 2015). Our results extend previous findings by showing invariant representation of subordinate object categories (car keys versus house keys). The differences in experimental design between our Experiment 2 and the approach used by Clarke and Tyler (2014) provide converging evidence for the representation of fine-

grained category information in PRC. These experimental designs can reveal complementary information, thus we propose that both approaches will be useful for future studies.

3.4.4 Conclusion

The context-guided object association task has played an instrumental role in advancing our understanding of representations within the MTL. Specifically, previous studies have shown that the hippocampus (Komorowski et al., 2009, 2013; Navawongse and Eichenbaum, 2013; Tort et al., 2013; McKenzie et al., 2014) and adjacent cortical regions (Keene et al., 2016) carry information about spatial and non-spatial aspects of events. Given that the MTL receives inputs from all sensory modalities, an interesting question is the degree to which representations in the MTL depend on low-level stimulus features. In Experiment 1, we provide evidence that the hippocampus, PHC, and RSC/PCC are related to the context matrix. In Experiment 2, we provide little evidence for invariant context representation, suggesting that the results from Experiment 1 were largely influenced by low-level stimulus features. Our results provide novel evidence for invariant object representation in PRC, which supports the notion that PRC is involved in the representation of fine-grained semantic information. Future studies should seek to elucidate the influence of physical versus psychological context on representations in MTL as well as potential differences between rodents and humans.

The results of our experiments provide valuable insight into the types of experiments that should be conducted to investigate item and context representation in the absence of low-level confounds. First, future studies could manipulate the context and items to a physically similar degree and within the same modality. Second, imaging/physiology data could be collected before and after training, including control stimuli which consist of identically structured events but without a learned behavioral repsonse or outcome. The inclusion of control stimuli is important because it can establish the role of associative memory. The

reverse type of experiment could also be used—i.e., testing whether representations of distinct contexts become more similar as a result of similar associations across the contexts. Third, the same stimuli can be used but with a different pre-event cue (e.g., a tone that signals the context, importantly this signal would terminate well before object sampling). Finally, future studies could further investigate invariant representation. The motivation for using multiple images in Experiment 2 was that any time that fixed stimuli are used to define a context or an object, there will necessarily be low-level differences between different contexts and objects. Importantly, previous research has revealed interesting results under conditions in which there were not observable low-level differences between stimuli (e.g., Marchette et al., 2015), similar to our results in PRC. It will also be interesting to examine the role of memory in invariant context and object representation.

Chapter 4

A behavioral and model-based investigation of human visual memory

Previous studies have indicated that healthy older adults are impaired in their ability to mnemonically discriminate between previously viewed stimuli and similar lure stimuli. We used a combined behavioral and model-based approach to test the hypothesis that healthy aging is accompanied by a deficit in the ability to encode stimulus features. For our behavioral approach, we used a forced-choice variant of a mnemonic similarity task. For our model-based approach, we found that a specific class of models from mathematical psychology, global matching models, provide a qualitatively good fit to our empirical data in both younger and healthy older adults. We found that decreasing the probability of successful feature encoding in the models resulted in a similar pattern of results to the empirical data. Collectively, our behavioral results extend to the forced-choice test format the finding that healthy aging is accompanied by an impaired ability to discriminate between targets and similar lures, and our modeling results suggest that a diminished probability of encoding stimulus features is a candidate mechanism for memory changes in healthy aging. We conclude by discussing the ability of global matching models to account for findings in other studies that have used mnemonic similarity tasks.

4.1   Introduction

Previous research has established that healthy older adults exhibit impaired performance on tests of associative memory. For example, a meta-analysis revealed that tests of source memory are impaired to a greater degree than tests of item memory (Spencer and Raz, 1995). More generally, Naveh-Benjamin and colleagues developed and tested an associative deficit hypothesis to account for memory changes among healthy older adults (e.g., Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2004; Old and Naveh-Benjamin, 2008b; for meta-analysis see: Old and Naveh-Benjamin, 2008a). Specifically, they reported a greater age-related impairment on tests of associative memory than tests of single item memory. Other studies have noted a greater impairment on memory recall tests than on traditional item recognition memory tests (i.e., targets vs unrelated foils; Craik and McDowd, 1987; Danckert and Craik, 2013). Taken together, there is unequivocal evidence for an age-related impairment on tasks that tax recollection and associative memory, with a more mild—and sometimes not statistically significant—deficit on tests of simpler item recognition memory.

Previous studies from our laboratory and others have shown that there are conditions in which healthy older adults exhibit a clear impairment on item recognition memory tasks. For example, our laboratory previously developed a Mnemonic Similarity Task, which assesses a participants' ability to discriminate between previously viewed objects (i.e., targets), similar lure objects, and unrelated foil objects (Kirwan and Stark, 2007; Yassa et al., 2011a,b; Kirwan et al., 2012; Stark et al., 2013, 2015). The ability to discriminate between targets and similar lures has been shown to be impaired in healthy older adults, with a relative sparing of their ability to discriminate between targets and unrelated foils (Tonor et al., 2009; Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015). Importantly, working memory versions of the task have failed to find age-related differences in the ability to discriminate between targets and similar lures (e.g., Yassa et al., 2011a), suggesting a mnemonic rather than a perceptual deficit in healthy older adults. Previous studies have used a test format in

which participants are instructed to respond "old" to exact repetitions of items seen during the encoding phase, to respond "similar" to images which are similar to—but not exactly the same as—a previously viewed image, and to respond "new" to images that they have not seen in the context of the experiment (Kirwan and Stark, 2007; Tonor et al., 2009; Yassa et al., 2011a,b; Kirwan et al., 2012; Stark et al., 2013, 2015; Bennett et al., 2015). Stark et al. (2015) also used a test format that instructed participants to respond "old" only to exact repetitions and to respond "new" to both similar lures and unrelated foils, including a version with confidence ratings. The results from these tests have consistently shown an age-related impairment in the ability to discriminate between targets and similar lures.

An unaddressed question is whether healthy older adults would exhibit impaired performance on a forced-choice variant of the Mnemonic Similarity Task. In the present study, we used several versions of the forced-choice procedure, similar to previous reports (Tulving, 1981; Holdstock et al., 2002; Migo et al., 2009, 2014; Jeneson et al., 2010). One benefit of the forced-choice test format is that it provides a method to assess memory performance which is free from the potential confound of between-group differences in response criterion (Green and Swets, 1966; Stanislaw and Todorov, 1999). In each of the test formats used here, we displayed one target object and one distractor object and participants were instructed to choose the exact object that they saw during the encoding phase; thus, we employed a two-alternative forced-choice procedure. In the first test format, participants were shown a target object and an unrelated foil, which we refer to as FC (A vs X), where FC is an acronym for forced-choice. In the second test format, participants were shown a target object and its corresponding similar lure, which we refer to as FCC (A vs A'), where FCC is an acronym for forced-choice corresponding (Migo et al., 2009). In the third test format, participants were shown a target object and a noncorresponding lure (i.e., a lure that is similar to a different object), which we refer to as FCNC (A vs B'), where FCNC is an acronym for forced-choice noncorresponding (Migo et al., 2009).

116

The dual-process complementary learning systems model has been used to advance the notion that patients with hippocampal damage will be impaired on the old/new test format with targets and similar lures and on the FCNC (A vs B') test format but will be relatively spared on the FCC (A vs A') test format, the FC (A vs X) test format, and the old/new test format with targets and unrelated foils (Norman and O'Reilly, 2003; see also: Holdstock et al., 2002; Migo et al., 2009, 2014). A study of a single patient with selective hippocampal damage revealed impaired performance on the old/new test format with targets and similar lures and intact performance on the FCC (A vs A') test format, supporting the predictions from the model (Holdstock et al., 2002). However, other studies with a larger sample of patients with selective hippocampal damage have shown a similar impairment on both the FCC (A vs A') test format and the old/new test format with targets and similar lures (Bayley et al., 2008; Jeneson et al., 2010). Additionally, Jeneson et al. (2010) revealed that the patients were equally impaired on the FCC (A vs A') test format, the FCNC (A vs B') test format, and the old/new test format with targets and similar lures. Although the results from patients with hippocampal damage are equivocal, the forced-choice test format can provide further insight into the organization of memory in younger adults and can elucidate the nature of memory changes that occur in the course of healthy aging.

Tulving (1981) used images of scenes to investigate performance across the three forced-choice test formats used in the present report. In a group of younger adults, Tulving (1981) revealed that there was an effect of test format, such that participants performed best on the FC (A vs X) test format, followed by the FCC (A vs A') test format, followed by the FCNC (A vs B') test format. Similarly, previous reports have shown better performance on the FCC (A vs A') test format than the FCNC (A vs B') test format in healthy middle-aged/older adults (Jeneson et al., 2010), in healthy older adults (Migo et al., 2014), and in patients with selective hippocampal damage (Jeneson et al., 2010). Hintzman (1988) revealed a similar superiority for the FCC (A vs A') test format compared to the FCNC (A vs B') test format in a study of word memory. Moreover, Hintzman (1988) showed that a global matching

model, MINERVA 2, could account for the effect of test format.

While previous studies have shown differences in performance based on the forced-choice test format, these studies have not investigated differences between younger adults and healthy older adults. Our primary aim was to investigate whether healthy older adults would exhibit impaired performance on the forced-choice version of the Mnemonic Similarity Task. Additionally, previous studies that have reported an effect of test format used images of scenes (Tulving, 1981), multiple encoding trials of images of objects (black and white silhouettes: Jeneson et al., 2010; Migo et al., 2014; color images: Jeneson et al., 2010), and judgments of the number of times that words were presented during the encoding phase (Hintzman, 1988). Our second aim was to investigate whether younger and older adults would exhibit an effect of test format in experiments that used images of objects (each target item of which was viewed once during the encoding phase). To address these questions, we conducted two behavioral experiments.

In Experiment 1, we included both younger and older adults and we used the three test formats mentioned above: FC (A vs X), FCC (A vs A'), FCNC (A vs B'). In Experiment 2, we aimed to replicate our findings from Experiment 1 and to rule out the possibility that the presence of the FC (A vs X) test format was artificially impairing performance on the FCNC (A vs B') test format. In a within-subjects design, participants performed two study-test cycles (with independent stimulus sets), one that included all three test formats and one that did not include the FC (A vs X) test format. Our third aim was to investigate whether a class of models from mathematical psychology—global matching models (e.g., Hintzman, 1984, 1988; Murdock, 1982, 1995)—could account for our empirical results in both younger and healthy older adults. Specifically, we tested the hypothesis that healthy aging could be modeled as an impaired ability to encode stimulus features. We conclude by discussing the application of global matching models to interpret the results of other experiments that have used the Mnemonic Similarity Task.

4.2 Materials and Methods

4.2.1 Experiment 1

Participants

Participants were 32 younger adults (18-28 years old) and 27 healthy older adults (64-85 years old). Older adults were screened to ensure that they did not have a memory impairment, similar to previous studies in our laboratory (Stark et al., 2013). Specifically, we ensured that participants scored in the normal range for their age group on the Mini-Mental Status Examination (MMSE; Crum et al., 1993) and the Rey Auditory Verbal Learning Task (RAVLT; Rey, 1941). We excluded 3 older adults because they did not score within 1.5 standard deviations of the mean for their age. Additionally, we excluded one younger adult due to very poor performance on our behavioral task (proportion correct ≈ 0.5 on all three test formats; more than 10 standard deviations below the mean of the included participants on the FC (A vs X) test format). Thus, 31 younger adults (26 female, 5 male) and 24 older adults (19 female, 5 male) were included in our analysis.

Behavioral tasks

Participants performed an incidental encoding task in which they indicated, via button press, whether they thought that the object in each picture was more of an "indoor" or an "outdoor" object (Fig. 4.1A; Stark et al., 2013, 2015). The encoding phase consisted of 140 images, which were displayed for 2000 ms with a 500 ms interstimulus interval. Following the encoding phase, participants performed a memory test, which contained three forced-choice test formats (top of Fig. 4.1B): 1) FC (A vs X; i.e., a target and an unrelated foil), 2) FCC (A vs A'; i.e., a target and a corresponding similar lure), 3) FCNC (A vs B'; i.e.,

a target and a noncorresponding lure [a lure object from a different pair]). On each test trial, one object was presented on the left side of the screen and one object was presented on the right side of the screen. Participants were told that on all trials they would view one image that they saw during the indoor/outdoor task and one new image. Moreover, they were told that on some trials the new image would be completely different than any of the images from the indoor/outdoor phase whereas on other trials the image would be similar to—but not exactly the same as—a previously viewed image from the indoor/outdoor phase. Participants were instructed to select, via button press, the exact image that they saw during the indoor/outdoor phase of the experiment. The images were displayed until the participant made a response or for 4 seconds, at which point the image disappeared and there was an unlimited response window. The target was randomly assigned to the left and right side of the screen on a trial-by-trial basis. The test formats were presented in a random, intermixed order. Participants performed 35 trials of each test format.

Our lab has previously calculated empirical estimates of the mnemonic similarity of the stimuli that we used in the present experiment (Lacy et al., 2011; Stark et al., 2013). Briefly, in over 100 participants, the mean proportion of times that participants responded "old" to a similar lure object was used as an index of mnemonic similarity (i.e., the higher the probability of responding "old" in response to a similar lure, the higher its mnemonic similarity). The stimuli were rank ordered and divided into 5 "lure bins." In the present experiment, we balanced the similarity of the stimuli at two levels: 1) the stimulus set: the number of trials from each lure bin in each test format (7 stimuli per lure bin), 2) the individual trial level: the lure bin of the target and distractor image. The former ensured that the similarity of targets and similar lures was balanced across the FCC (A vs A') and the FCNC (A vs B') test formats for every subject. The latter addressed the potential issue of encoding vs retrieval difficulty of stimuli from different lure bins, which is particularly important for the FCNC (A vs B') test format.

Behavioral data analysis

We calculated the proportion correct for each test format for each participant. To investigate whether there was an effect of test format, irrespective of age, we performed a separate one-way analysis of variance (ANOVA) in each age group. We performed planned tests to investigate whether performance was ranked in the following order: 1) FC (A vs X), 2) FCC (A vs A'), 3) FCNC (A vs B'). To investigate whether there was an age by test format interaction, we performed a mixed-design ANOVA (between-subjects variable: Age Group, within-subjects variable: Test Format). We performed planned tests to investigate whether performance differed between younger and older adults on the FCC (A vs A') test format and on the FCNC (A vs B') test format.

4.2.2   Experiment 2

Participants

Participants were 21 younger adults (18-33 years old). We excluded one participant due to very poor performance on our behavioral task (proportion correct ≈ 0.5 on all of the test formats; more than 10 standard deviations below the mean of the included participants on the FC (A vs X) test format). Thus, 20 participants (14 female, 6 male) were included in our analysis.

Behavioral tasks

The behavioral tasks in Experiment 2 were similar to Experiment 1. In Experiment 2, participants performed two encoding and two testing phases, each with a distinct stimulus set. Previous research in our laboratory has ensured that the two stimulus sets are perfectly

matched in terms of similarity (Stark et al., 2015). The encoding phases were identical to those in Experiment 1. The two test phases differed in the number of test formats used. The purpose of this manipulation was to address whether the FC (A vs X) test format was artificially reducing performance on the FCNC (A vs B') test format. Accordingly, one version used three test formats, as in Experiment 1, and the other version used two test formats: 1) FCC (A vs A'), 2) FCNC (A vs B'). In the two test version, participants were instructed that on each trial they would view one image that was in the indoor/outdoor task and one image that was similar to—but not exactly the same as—an image from the indoor/outdoor task. Moreover, they were instructed that on some trials the similar image would be from the same pair (e.g., if they studied an image of an apple they might see the exact apple and a similar apple) and on some trials the similar image would be from a different pair (e.g., if they studied an apple and an orange, they might see the exact apple and a similar orange). The order in which participants received the three and the two test format condition was counterbalanced between participants. As in Experiment 1, participants performed 35 trials of each test format. Thus, the test phase contained 35 fewer trials in the two test condition.

Behavioral data analysis

A one-way ANOVA was used to test the presence of a main effect of test format in the three test version. We performed planned tests to investigate whether performance was ranked in the following order: 1) FC (A vs X), 2) FCC (A vs A'), 3) FCNC (A vs B'). For the two test version, we performed a planned test to investigate whether performance was better on the FCC (A vs A') test format than the FCNC (A vs B') test format. Additionally, we investigated whether performance was enhanced on the two test format relative to the three test format using separate paired t-tests for the FCC (A vs A') and the FCNC (A vs B') test formats.

4.2.3 Global matching models

MINERVA 2

MINERVA 2 (Hintzman, 1984, 1988) is a member of a class of mathematical psychology models referred to as global matching models. MINERVA 2 is a multiple-trace or exemplar-based model, meaning that a new memory trace is added to an existing memory matrix every time that an item is encoded. In MINERVA 2, items are represented as vectors, each feature of which is set to -1, 0, or 1 with equal probability (i.e., 1/3). Similar lures were generated for each target by re-drawing from the original features with probability δ. In the present report, we used $\delta = 0.16$, meaning that on average 16% of the features were re-drawn from the original distributions. This resulted in approximately 11% of the features changing values. During encoding, each feature is encoded with probability of L and not encoded with probability $1 - L$. The encoding phase results in a memory matrix, $\mathbf{T}$, which contains M rows (i.e., memory traces) and N columns (i.e., features). Our implementation relied on the equations presented in (Hintzman, 1984, 1988) and our simulations used $M = 35$, similar to our empirical test formats. The first equation provides an estimate of the similarity of a probe ($\mathbf{p}$; i.e., a test item) to a given trace ($\mathbf{T}_i$; i.e., one of the items in memory):

$$\mathbf{s}_i = \left(\frac{\mathbf{p} \cdot \mathbf{T}_i}{n_i} \right)^3 \tag{4.1}$$

where n_i is the number of features that are relevant to the comparison of the probe and a given trace (a feature is relevant if it is non-zero in either $\mathbf{p}$ or $\mathbf{T}_i$). Thus, the portion of the equation within the parentheses is a normalized dot product. The cubing function causes the similarity function to be nonlinear, which allows retrieval to be "quite selective" (Hintzman, 1984, 1988). The global match, g, of the trace is given by the summed similarity across all

stored traces (where there are M traces in the memory matrix):

$$g = \sum_{i=1}^{M} \mathbf{s}_i \tag{4.2}$$

While MINERVA 2 uses a multiple-trace storage operation, the retrieval operation is the global match of a probe to all of the contents in memory. Thus, MINERVA 2 is a global matching model by the nature of its retrieval process. We modeled MINERVA 2 in R.

TODAM

TODAM (Theory of Distributed Associative Memory; Murdock, 1982) is a different global matching model. In contradistinction to MINERVA 2, which is a multiple-trace or exemplar-based model, TODAM is a distributed or prototype-based memory model, meaning that memories are stored in a single, composite memory vector (e.g., a prototype). Thus, while these models share the assumption that memory retrieval is a global matching process, the memory storage mechanisms of the models are very different. While most versions of TODAM have focused on associative memory tasks (e.g., Murdock, 1982), it can also be used as an item-only model (e.g., Murdock, 1995). Our implementation relied on the version of TODAM presented in (Murdock, 1995).

As in MINERVA 2, items are represented as vectors. In TODAM, each feature of an item vector is a random draw from a normal distribution with mean 0 and standard deviation $\sqrt{1/N}$, where N is the number of features. Occasionally the numerator is set to a value other than 1 (this parameter is referred to as P in Murdock, 1982); however, setting the value to 1 causes the vectors to be of approximately unit length which is useful because similarity is calculated using the dot product (i.e., the dot product between two vectors of unit length is between -1 and 1, similar to a normalized dot product). The following equation was used to

generate a similar lure item ($\mathbf{f}_{j'}$) for a given target item ($\mathbf{f}_j$; Murdock, 1995):

$$\mathbf{f}_{j'} = \rho\mathbf{f}_j + \left(\sqrt{1 - \rho^2}\right)\mathbf{g}_j \tag{4.3}$$

where ρ represents the similarity of items to each other and $\mathbf{g}_j$ represents an independent random vector. The expected value of the similarity, defined as the dot product, between a target item and its similar lure is ρ. The memory vector for the item-only version of TODAM was calculated with the following equation (Kahana, 2012, pg. 105; Murdock, 1995):

$$\mathbf{m}_t = \alpha\mathbf{m}_{t-1} + \mathbf{B}_t\mathbf{f}_t \tag{4.4}$$

where α is a forgetting parameter (which can also be thought of as a retention parameter because 0 represents complete erasure of previous memories whereas 1 indicates that the new memory is added to the memory vector from the previous trial without any forgetting), $\mathbf{m}_{t-1}$ represents the memory vector from the previous trial, and $\mathbf{f}_t$ represents the item that is presented at time t in the encoding phase. $\mathbf{B}_t$ is a diagonal matrix with entries drawn from a Bernoulli distribution with probability p, where p represents the probability that a feature is encoded (Kahana, 2012, pg. 105)—i.e., each feature is encoded with probability p and not encoded with probability $1 - p$ (Murdock, 1995). Accordingly, p is isomorphic to L in MINERVA 2. As implied by the subscript t, $\mathbf{B}_t$ is trial unique. For item memory, the model has four parameters: 1) α, the forgetting/retention rate, 2) N, the number of features in each item, 3) p, the probability of encoding a feature, and 4) ρ, the similarity between a target item and its lure. To preserve similarity to MINERVA 2, α was set to 1. Thus, in our application, both models have 3 parameters: 1) the number of features, 2) the probability of encoding a feature, 3) the similarity between targets and similar lures. Additionally, we used a list length of 35 items as in our MINERVA 2 simulations and in our empirical test formats. The global match, g, of a probe to the memory vector was calculated with the

following equation:

$$g = \mathbf{p} \cdot \mathbf{m} \tag{4.5}$$

where $\mathbf{p}$ represents a probe item, and $\mathbf{m}$ represents the memory vector. In contrast to MIN-ERVA 2 (Equation 4.1), the item-only version of TODAM uses a linear similarity function. Also, because TODAM uses a single, composite memory vector, the global match is simply defined as the similarity—i.e., the dot product—between the probe and the memory vector. Thus, TODAM is a global matching model by the nature of both its storage and its retrieval operations. The standard instantiations of TODAM use closed-form equations to calculate measures such as d'; however, we were interested in the effect of test format on forced-choice performance. Thus, we used a computational, rather than a mathematical, approach. We modeled TODAM in GNU Octave.

Simulation of forced-choice performance

As in our empirical data, we were interested in simulating performance from three different test formats: 1) FC (A vs X), 2) FCC (A vs A'), 3) FCNC (A vs B'). To simulate the FC (A vs X) test format, we calculated the proportion of times that the global match, g (Equations 4.2 and 4.5), for a target item (A) exceeded that of an unrelated foil (X), using the following equation (cf. Hintzman, 1988):

$$Pr\left\{A > X\right\} = \frac{1}{M} \sum_{i=1}^{M} [I(g_{A_i} > g_{X_i}) + 0.5 \cdot I(g_{A_i} = g_{X_i})] \tag{4.6}$$

where M is the list length and $I(\cdot)$ is the indicator function which sets the value to 1 if the statement is true and to 0 otherwise. The second part of the equation simulates random guessing if the two items generate the same global match. To simulate the FCC (A vs A') test format, we calculated the proportion of times that the global match for a target item (A)

exceeded that of its lure item (A') using Equation 4.6. Similarly, to simulate the FCNC (A vs B') test format, we calculated the proportion of times that the global match for a target item (A) exceeded that of a similar lure item from a different pair (B') using Equation 4.6. In both models, we simulated 10,000 participants and we determined parameter values that provided a good fit to the empirical data for the younger adults. To test the hypothesis that healthy aging is accompanied by impaired encoding, we investigated the effect of decreasing the encoding parameter in both models (L and p in MINERVA 2 and TODAM, respectively).

4.3 Results

4.3.1 Experiment 1

We first investigated whether there was an effect of test format on performance in both age groups. Separate one-way ANOVAs revealed a significant main effect of test format in both younger adults ($F = 110.7$, $p < 0.001$) and older adults ($F = 100.2$, $p < 0.001$). Planned comparisons revealed that both age groups performed better on the FC (A vs X) test format than the FCC (A vs A') test format (YA: $t_{30} = 8.97$, $p < 0.001$; OA: $t_{23} = 11.36$, $p < 0.001$) and better on the FCC (A vs A') test format than the FCNC (A vs B') test format (YA: $t_{30} = 6.63$, $p < 0.001$; OA: $t_{23} = 3.58$, $p < 0.005$). These results suggest that there is an effect of test format in both age groups. Notably, although performance was the worst on the FCNC (A vs B') test format, performance was significantly better than chance in both age groups (YA: $t_{30} = 9.33$, $p < 0.001$; OA: $t_{23} = 9.13$, $p < 0.001$).

We next investigated whether there was an effect of healthy aging on performance. A mixed-design ANOVA (between-subjects variable: age group, within-subjects variable: test format) revealed a significant age group by test format interaction ($F = 3.51$, $p = 0.033$). Planned comparisons revealed that younger adults performed significantly better than older adults
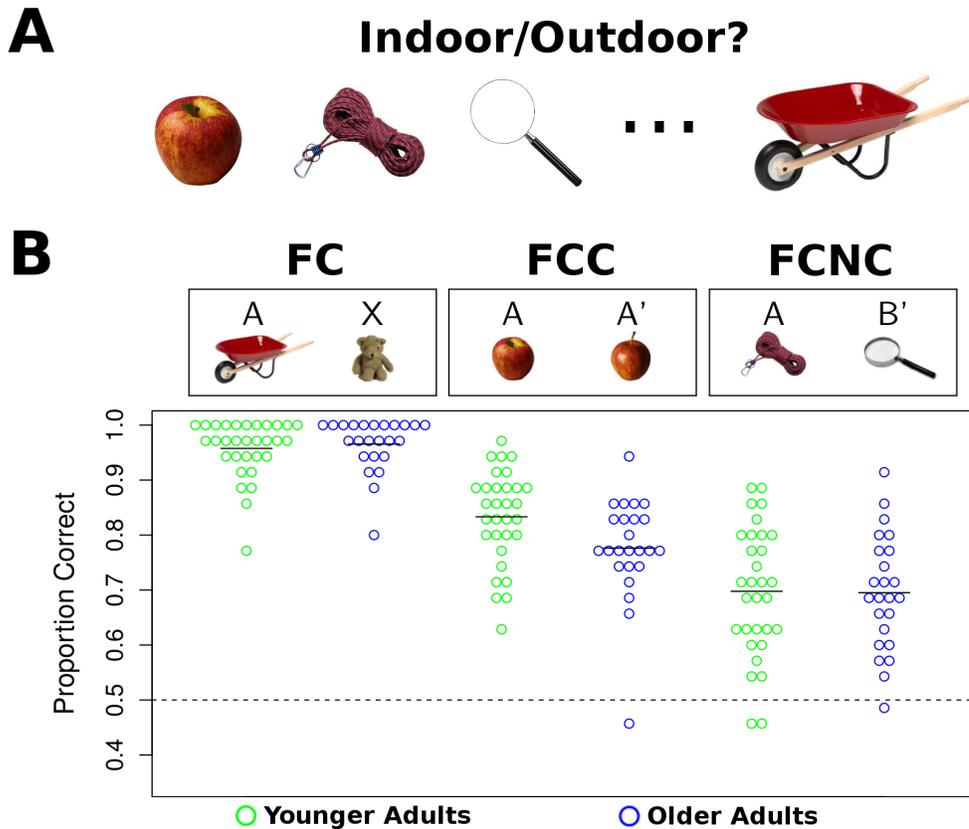
Figure 4.1: Forced-choice performance. A) The encoding task was an incidental design in which participants indicated whether each item was an indoor or an outdoor item. B) The test phase consisted of three test formats (FC: A vs X; FCC: A vs A'; FCNC: A vs B'). In both age groups, there was a significant effect of test format. A repeated-measures ANOVA revealed an age by test format interaction, which was driven by better performance of younger adults on the FCC (A vs A') test format. Circles represent individual participants and lines represent the mean proportion correct.

on the FCC (A vs A') test format ($t_{53} = 2.37$, $p < 0.025$). Conversely, the difference between younger and older adults failed to reach significance for the FCNC (A vs B') test format ($t_{53} = 0.081$, $p = 0.94$; Fig. 4.1). These results extend the previous findings of an age-related decline in performance on the old/new and the old/similar/new test format with targets and similar lures (Tonor et al., 2009; Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015) to the FCC (A vs A') test format.

Previous studies that used the old/similar/new test format reported an age-related impair-

ment in the ability to discriminate between targets and similar lures with intact discrimination between targets and unrelated foils (Tonor et al., 2009; Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015). Similarly, there was no evidence to suggest that younger adults performed better than healthy older adults on the FC (A vs X) test format ($t_{53} = -0.58$, $p = 0.57$). It is possible that an age-related difference on the FC (A vs X) test format was obscured by a ceiling effect; however, when we compared the 15 worst performing younger adults (i.e., median split) and the 12 worst performing older adults (i.e., median split), the difference still failed to reach significance ($t_{25} = -0.56$, $p = 0.58$). Taken together, the results provide clear evidence for an effect of test format in both age groups and the results suggest that healthy older adults are impaired at discriminating between targets and their similar lures—i.e., the FCC (A vs A') test format.

4.3.2 Experiment 2

Main results

In Experiment 2, we aimed to replicate the findings in younger adults from Experiment 1 as well as to rule out the possibility that the FC (A vs X) test format artificially reduced FCNC (A vs B') test format performance. Specifically, we thought that it was possible that the FC (A vs X) test format increased the propensity for participants to immediately select the first item that they viewed in the FCNC (A vs B') test format. Thus, in Experiment 2, participants performed two study-test cycles (with distinct stimulus sets), one of which included all three test formats and the other of which included only the FCC (A vs A') test format and the FCNC (A vs B') test format.

In the three test condition, a one-way ANOVA revealed a significant main effect of test format ($F = 73.1$, $p < 0.001$; Fig. 4.2A). Planned comparisons revealed significantly better performance on the FC (A vs X) test format than the FCC (A vs A') test format ($t_{19} = 9.36$,
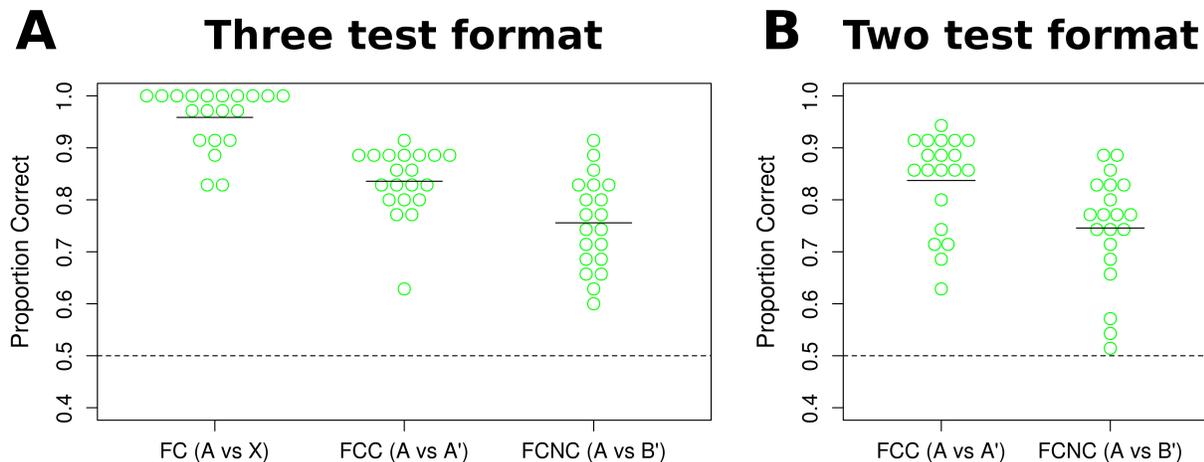
Figure 4.2: Investigation of performance on the three and two test conditions in younger adults. A) A one-way ANOVA revealed a significant effect of test format on performance ($F = 73.1$, $p < 0.001$). Paired t-tests revealed significantly better performance on the FCC (A vs A') condition compared to the FCNC (A vs B') condition for both the three test condition ($t = 4.38$, $p < 0.001$) and the two test condition ($t = 3.05$, $p < 0.01$; see Fig. B). Paired t-tests revealed no sign of a benefit for the two test format over the three test format for either the FCC (A vs A') format ($M = 0.0014$, $t_{19} = 0.078$, $p = 0.94$) or the FCNC (A vs B') format ($M = -0.01$, $t_{19} = -0.46$, $p = 0.65$).

$p < 0.001$) and significantly better performance on the FCC (A vs A') test format than the FCNC (A vs B') test format ($t_{19} = 4.38$, $p < 0.001$). In the two test condition, a paired t-test revealed significantly better performance on the FCC (A vs A') test format than the FCNC (A vs B') test format ($t_{19} = 3.05$, $p < 0.01$; Fig. 4.2B). Finally, paired t-tests revealed no sign of a benefit for the two test format over the three test format for either the FCC (A vs A') test format ($M = 0.0014$, $t_{19} = 0.078$, $p = 0.94$) or the FCNC (A vs B') test format ($M = -0.01$, $t_{19} = -0.46$, $p = 0.65$). These results replicate the effect of test format that we observed in Experiment 1. Moreover, these results rule out the possibility that the FC (A vs X) test format was artificially reducing performance on the FCNC (A vs B') test format.

Comparison of performance on forced-choice and old/new test formats

The results of Experiment 1 and 2 provide clear evidence to suggest that performance is better on the FCC (A vs A') test format than the FCNC (A vs B') test format. Previous

130

reports have suggested that performance on the FCC (A vs A') test format can rely on familiarity to a greater degree than performance on the FCNC (A vs B') test format and the old/new test format with targets and similar lures (Holdstock et al., 2002; Migo et al., 2009, 2014). Accordingly, we were interested in examining whether performance on the FCC (A vs A') test format was better than performance on the old/new test format with targets and similar lures. Similarly, we were interested in investigating whether performance was better on the FC (A vs X) test format than performance on the old/new test format with targets and unrelated foils.

To compare old/new performance with performance on the forced-choice tests from Experiment 2, we reanalyzed data from 20 younger adults from a previous study from our laboratory (Experiment 4 in Stark et al., 2015). As in Experiment 2, participants performed two study-test cycles with two unique stimulus sets. The encoding phase consisted of an indoor/outdoor judgment for each of 128 images of objects. One of the test formats used "gist" instructions (i.e., participants were instructed to respond "old" to similar lures) while the other test format used "veridical" instructions (i.e., participants were instructed to respond "new" to similar lures), and the order of the test formats was counterbalanced across participants. For the present analysis, we used the data from the veridical condition because the test instructions were equivalent to our instructions for Experiment 2. The test phase consisted of three probe types: 1) targets (exact repetitions), 2) similar lures, 3) unrelated foils. Participants were instructed to respond "old" only for exact repetitions and to respond "new" for both similar lures and for novel foils. After making the old/new decision, participants indicated the confidence of their response (very sure, somewhat sure, somewhat unsure, very unsure), resulting in 8 confidence bins (ranging from "very sure old" to "very sure new"). Participants performed 64 trials of each probe type. Three participants were excluded due to a failure to distribute responses across the confidence bins (which resulted in poor model fit), thus 17 participants were included in the between-group analysis.

131
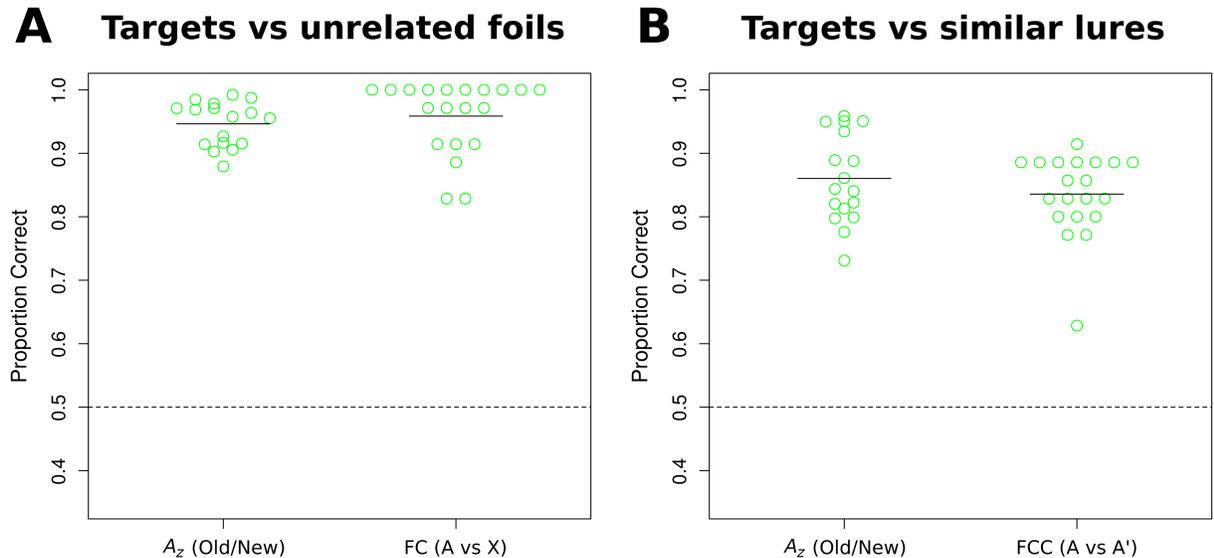
Figure 4.3: Between-group performance was similar on forced-choice and old/new test formats. A) A_z and two-alternative forced-choice performance were similar for targets vs unrelated foils ($t_{35} = -0.76$, $p = 0.45$). B) A_z and two-alternative choice performance were similar for targets vs similar lures ($t_{35} = 1.11$, $p = 0.28$; see Fig. B). In contrast, A_z for targets vs similar lures was significantly greater than performance on the FCNC (A vs B') test format ($t_{35} = 3.96$, $p < 0.001$).

The area under the receiver operating characteristic (ROC) curve—calculated from the old/new test format with confidence ratings—is mathematically equivalent to proportion correct on the two-alternative forced-choice test format (Green and Moses, 1966; Green and Swets, 1966; Swets and Pickett, 1982; Stanislaw and Todorov, 1999). The preferred approach for estimating the area under the ROC curve is to use maximum-likelihood estimation to fit the z-transformed ROC curve—a measure referred to as A_z (Swets and Pickett, 1982; Stanislaw and Todorov, 1999). Importantly, A_z does not assume equal variance of the target and distractor (e.g., unrelated foil, similar lure) distributions. We used the function *rocfit* in Stata to compute A_z.

If the FCC (A vs A') test format enhances a participants ability to rely on familiarity, then we should observe significantly better performance on the FCC (A vs A') test format (i.e., proportion correct) than on the old/new test format (i.e., A_z). Conversely, if performance on the FCC (A vs A') test format and the old/new test format rely on similar cognitive

processes, then we should not observe a difference between proportion correct on the FCC (A vs A') format and A_z from the old/new format. Similarly, if the FC (A vs X) test format enhances a participants ability to rely on familiarity, then we should observe significantly better performance on the FC (A vs X) test format than on the old/new test format.

The difference between A_z for targets vs unrelated foils and proportion correct on the FC (A vs X) test format failed to reach significance ($t_{35} = -0.76$, $p = 0.45$; Fig. 4.3A). Additionally, the difference between A_z for targets vs similar lures and proportion correct on the FCC (A vs A') test format failed to reach significance ($t_{35} = 1.11$, $p = 0.28$; Fig. 4.3B). Taken together, these results suggest that the old/new test format and the forced-choice test format recruit similar cognitive processes. In contrast, A_z for targets vs similar lures was significantly greater than proportion correct on the FCNC (A vs B') test format ($t_{35} = 3.96$, $p < 0.001$). While there were minor differences between the stimulus sets used in these two experiments, these results are at least consistent with the notion that that the old/new test format with targets vs similar lures is more closely related to the FCC (A vs A') test format. Finally, for comparison to the present experiments, a paired t-test revealed significantly greater A_z for targets vs unrelated foils than A_z for targets vs similar lures ($t_{16} = 6.93$, $p < 0.001$).

Summary of the behavioral experiments

In Experiment 1, we found that younger adults performed better than healthy older adults on the FCC (A vs A') test format, similar to previous studies that used the old/similar/new test format with targets and similar lures (Tonor et al., 2009; Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015). In Experiment 1, we provided clear evidence for an effect of test format in both age groups, such that participants performed best on the FC (A vs X) test format, followed by the FCC (A vs A') test format, followed by the FCNC (A vs B') test format. In Experiment 2, we replicated the effect of test format in a group of younger adults. Moreover, we ruled out the possibility that impaired performance on the FCNC (A

vs B') test format was driven by the presence of the FC (A vs X) test format. The finding that performance on the FCNC (A vs B') test format was reliably worse than performance on the FCC (A vs A') test format was initially puzzling; however, previous reports have shown that global matching models can account for this effect (Hintzman, 1988, 2001, also see: Clark and Gronlund, 1996). Accordingly, we were interested in investigating whether global matching models could be used to account for our results in both younger and healthy older adults.

4.3.3 Global matching models

MINERVA 2

We investigated whether MINERVA 2 could account for our empirical findings of Experiment 1 in both younger and older adults. Specifically, we tested the hypothesis that aging could be modeled as a decreased probability of encoding stimulus features. We began by finding model parameters that achieved similar values to the mean values of our empirical data in younger adults ($N = 20$, $L = 0.65$, and $\delta = 0.16$; Fig. 4.4A). We next investigated whether decreasing the encoding parameter, L, would cause a similar pattern of deficits as we observed in healthy older adults. Decreasing the encoding parameter from $L = 0.65$ to $L = 0.55$ resulted in worse performance on all 3 test formats; however, the largest change in performance was on the FCC (A vs A') format, which was the format with a significant age group difference in the empirical data (Fig. 4.4B). Thus, at least for certain model parameters, MINERVA 2 can predict a disproportionate change in performance on the FCC (A vs A') test format by simply changing the L parameter. These results support the hypothesis that healthy aging is accompanied by an impaired ability to encode stimulus features. We next asked whether a different global matching model, which relies on a very different storage mechanism, would predict a similar pattern of results.
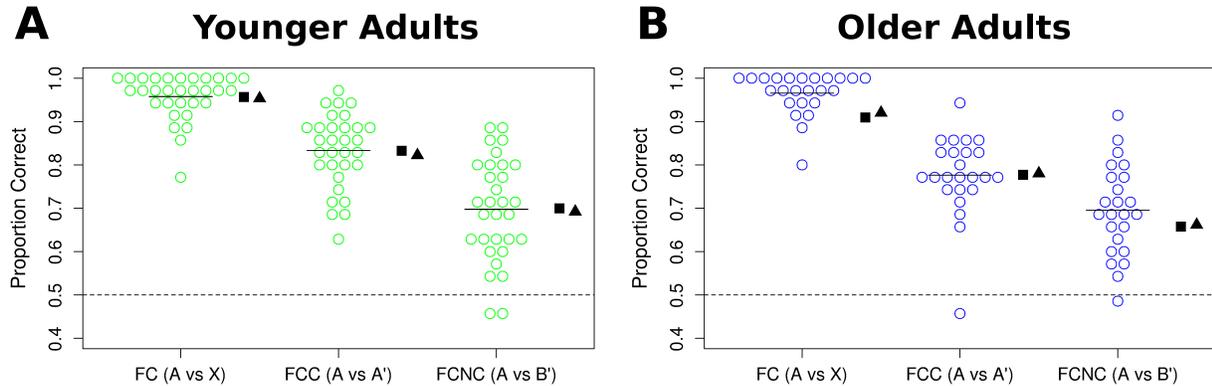
Figure 4.4: MINERVA 2 (squares) and TODAM (triangles) were both able to capture the main pattern of results of the empirical data. Specifically, both models captured the FC (A vs X) > FCC (A vs A') > FCNC (A vs B') effect, and decreasing the encoding parameter in both models caused the largest change in performance on the FCC (A vs A') test format. A) Empirical data and model fit data for younger adults/$L = 0.65$/$p = 0.5$. B) Older adults/$L = 0.55$/$p = 0.35$. Black squares represent the mean of the MINERVA 2 data and black triangles represent the mean of the TODAM data.

TODAM

As in the MINERVA 2 simulation, we began by finding parameters that achieved similar values to the mean values of our empirical data in younger adults ($N = 400$, $p = 0.5$, and $\rho = 0.7$; Fig. 4.4A). We next investigated whether decreasing the encoding parameter, p, would cause a similar pattern of deficits as we observed in healthy older adults. Decreasing the encoding parameter from $p = 0.5$ to $p = 0.35$ resulted in worse performance on all 3 test formats; however, the largest change was on the FCC (A vs A') format, which was the format with a significant age group difference in the empirical data (Fig. 4.4B). Thus, at least for certain model parameters, TODAM can predict a disproportionate change in performance on the FCC (A vs A') test format by simply changing the p parameter. Given that MINERVA 2 and TODAM use very different storage mechanisms, these results provide additional support for both the global matching framework and for the hypothesis that healthy aging is accompanied by an impaired ability to encode item features.

Why do the models predict better performance on the FCC (A vs A') test format than the FCNC (A vs B') test format?

Previous reports showed that both MINERVA 2 and TODAM predict better performance on the FCC (A vs A') test format relative to the FCNC (A vs B') test format (Hintzman, 1988, 2001; Clark and Gronlund, 1996). As discussed by Hintzman (1988, 2001), variability increases the overlap between target and distractor (e.g., similar lure, unrelated foil) distributions in MINERVA 2. One source of variability in MINERVA 2 (and TODAM) is encoding variability. In the standard version of MINERVA 2, each feature is encoded with probability L; hence, on average $L \times N \times 2/3$ non-zero features are encoded for each item, where N is the total number of features and it is multiplied by $2/3$ because on average one third of the features are equal to zero. Because the number of encoded non-zero features is variable, there are trials where the number of non-zero features that are encoded is greater than $L \times N \times 2/3$ and trials where the number of non-zero features that are encoded is less than $L \times N \times 2/3$.

We hypothesized that removing trial-by-trial encoding variability in MINERVA 2 would reduce the FCC (A vs A') test format advantage. We tested this hypothesis by altering the model to encode a fixed number of features on each trial (note, a similar approach would be more difficult in TODAM because the features are drawn from a normal distribution rather than from {-1, 0, 1}). First, we set the number of non-zero features to be equal on each trial. In this version of the model we increased N from 20 to 21 to allow an equal number of -1, 0, and 1 features (i.e., 7 each) and we set $L = 9/14$. We verified that this had no effect on performance of the model (see "with encoding variability" in Fig. 4.5). Next, we eliminated encoding variability by forcing the model to encode 9 of the 14 non-zero features. Thus, the only difference between these two models is the presence of encoding variability. We observed an increase in proportion correct for all formats, and we observed a reduction of the FCC (A vs A') test format advantage over the FCNC (A vs B') test format (Fig. 4.5).

Effect of encoding variability
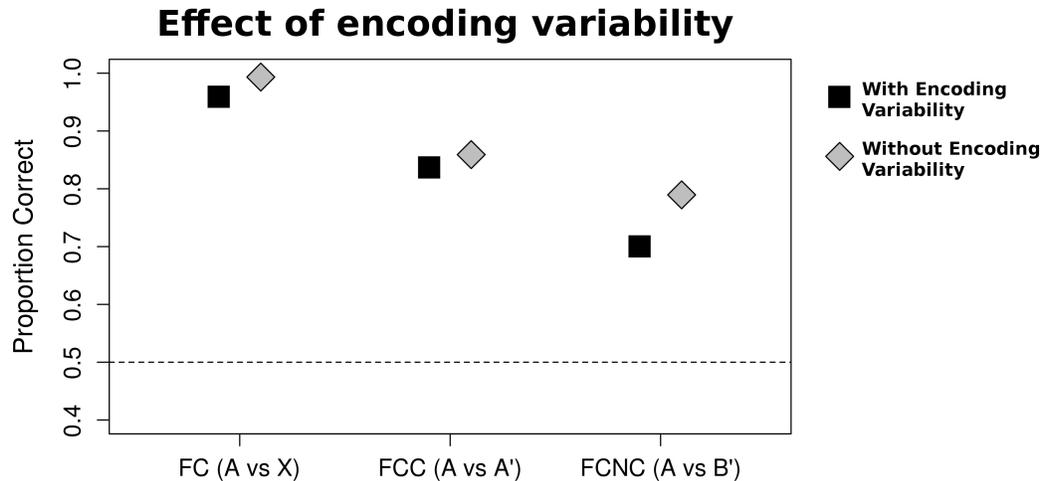
Figure 4.5: The removal of encoding variability in MINERVA 2 reduced the magnitude of the difference between proportion correct on the FCC (A vs A') format and the FCNC (A vs B') format. Data points represent the mean proportion correct.

These results suggest that one possible reason for worse performance on the FCNC (A vs B') test format relative to the FCC (A vs A') test format is that for some trials participants happen to encode more features for the original B item than the original A item. Because the lures are correlated with the original target item, this results in greater summed similarity for the B' item than the A item. Under the condition in which there is not variability in the number of features that are encoded for each A and B item, there is less of a difference in performance between the FCC (A vs A') test format and the FCNC (A vs B') test format.

Encoding variability reduced the FCC (A vs A') test format advantage but it did not eliminate the advantage. As discussed by Hintzman (1988, 2001), there are other sources of variability that contribute to the FCC (A vs A') test format advantage. For example, within the stimulus set used for the encoding phase, some stimuli happen to be more similar to other stimuli, which results in certain trials in which the B item more closely resembles other items in the encoding set than the A item. Because the model assumes that memory strength is determined by the match of the test item to all of the contents of memory, this results in a greater global match of the B' item than the A item (Hintzman, 1988, 2001). Interestingly, in our simulations, we found that list length modulated the strength of the

effect of encoding variability on the FCC (A vs A') test format advantage. Specifically, for shorter list lengths (e.g., 4 items), the elimination of encoding variability accounted for more of the difference between the two test formats than for longer list lengths (e.g., 35 items). In fact, for short list lengths, the elimination of encoding variability was sufficient to nearly eliminate the difference between the two test formats, suggesting that as more items are encoded there is a greater chance of a B' item providing a better global match than the A item (due to similarity to other items in the stimulus set). Thus, the models suggest that there are a number of potential sources of variability that contribute to enhanced FCC (A vs A') test format performance compared to the FCNC (A vs B') test format, including encoding variability and the similarity between items in the study list.

4.4 Discussion

4.4.1 The effect of test format on performance

We investigated the effect of test format on recognition memory performance in younger and healthy older adults. In Experiment 1, we used three test formats: 1) FC (A vs X), 2) FCC (A vs A'), 3) FCNC (A vs B'). There was a significant effect of test format in both younger and older adults. Specifically, in both age groups, performance was best on the FC (A vs X) format, followed by the FCC (A vs A') format, followed by the FCNC (A vs B') format. The results from Experiment 2 replicated the results of Experiment 1 and provided no evidence to suggest that the FC (A vs X) test format artificially reduced performance on the FCNC (A vs B') test format. Taken together, we consistently showed significantly better performance on the FCC (A vs A') test format than the FCNC (A vs B') test format. The findings in younger adults replicate the findings from Tulving (1981). Similarly, other reports have shown enhanced performance on the FCC (A vs A') test format compared

138

to the FCNC (A vs B') test format in young adults (Hintzman, 1988; but see: Migo et al., 2009), healthy middle-aged/older adults (mean age: 61.2 years; Jeneson et al., 2010), healthy older adults (mean age 71; Migo et al., 2014), and in patients with selective hippocampal damage (Jeneson et al., 2010). Moreover, performance on the FCC (A vs A') test format is consistently better than performance on the FCNC (A vs B') test format across a variety of encoding and stimulus conditions: single presentations of images of objects (Experiment 1 and Experiment 2) and of scenes (Tulving, 1981), multiple encoding trials of images of objects (color images: Jeneson et al., 2010; black and white silhouettes: Jeneson et al., 2010; Migo et al., 2014; but see: Migo et al., 2009), and judgments of the number of times that words were presented during the encoding phase (Hintzman, 1988).

Similar to Hintzman (1988, 2001), we showed that MINERVA 2 could account for the observed effect of test format. Moreover, we showed that a different global matching model, TODAM (Murdock, 1982, 1995), can also account for the observed effect of test format (cf. Clark and Gronlund, 1996). We used MINERVA 2 to provide a possible explanation of the FCC (A vs A') test format advantage (also see: Hintzman, 1988, 2001). Specifically, we showed that removing trial-by-trial encoding variability reduced the magnitude of the FCC (A vs A') test format advantage. Thus, the model suggests that one possible reason for the FCC (A vs A') test format advantage is that there are trials on which a participant happens to encode more details than other trials, which causes certain lures (B') to contain a stronger global match than a noncorresponding target item (A)—i.e., the global match for the lures is shifted along with the global match of the target item due to the similarity between them. Additionally, we found that the list length contributed to the effect of encoding variability, such that the elimination of encoding variability had a larger effect for short list lengths. Specifically, for short lists, the removal of encoding variability nearly eliminated the FCC (A vs A') test format advantage. Our simulations with longer list lengths suggested that there are more trials in which the B item (and by extension the B' item) is more similar to other items in the encoding set than the A item. Accordingly, both encoding variability

139

and variability in between-item similarity in the encoding list could contribute to better performance on the FCC (A vs A') test format than the FCNC (A vs B') test format.

4.4.2 Forced-choice and old/new test formats reveal a stable age-related impairment of performance

We next investigated whether the previous reports of an age-related impairment on the Mnemonic Similarity Task (Tonor et al., 2009; Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015) would extend to the forced-choice test format. Our results revealed a significant age by test format interaction, which was driven by better performance in younger adults than healthy older adults on the FCC (A vs A') test format. These results suggest that the age-related impairment on the old/similar/new and the old/new test formats with targets and similar lures extends to the FCC (A vs A') test format. Similarly, in a group of healthy older adults, Migo et al. (2014) revealed a significant relationship between age and proportion correct on the FCC (A vs A') test format but not between age and proportion correct on the FCNC (A vs B') test format. We have previously reported a relationship between age and mnemonic discrimination performance in a lifespan sample that used the old/similar/new test format (Stark et al., 2013; Bennett et al., 2015). Taken together, these findings support the notion that the old/similar/new and FCC (A vs A') test formats are similarly affected by aging. We were admittedly surprised that the age difference in performance on the FCNC (A vs B') test format failed to reach significance; however, we suggest that the results from our model-based approach and the results from Experiment 2 provide possible explanations for the significant age-related difference on the FCC (A vs A') test format but not on the FCNC (A vs B') test format, which we will discuss in turn.

We investigated whether global matching models could account for the empirical changes observed in healthy older adults. As mentioned above, previous studies have shown that global

matching models can account for the FCC (A vs A') test format advantage compared to the FCNC (A vs B') test format. Global matching models have few parameters; accordingly, the results of simulations from global matching models are straightforward to interpret. We started by finding parameters in the models that provided a good fit of the empirical data in younger adults. We next investigated the effect of decreasing the probability of encoding stimulus features. In MINERVA 2 and TODAM, decreasing the encoding probability caused the largest change in performance on the FCC (A vs A') test format, which was the test format on which we observed an age-related change. It is noteworthy, however, that both models predicted a change on the other test formats as well, which suggests that the most sensitive test format for detecting differences in encoding was the FCC (A vs A') test format. MINERVA 2 and TODAM rely on a very different set of assumptions regarding how memories are stored—namely, MINERVA 2 is a multiple-trace model while TODAM is a distributed memory model. The fact that both models predicted the largest change on the FCC (A vs A') test format as a result of decreasing the encoding parameter supports the global matching framework and suggests that a possible explanation for the observed age-related changes is a decrease in the probability of encoding stimulus features.

4.4.3 Younger adults perform similarly on old/new and forced-choice test formats

Performance on two-alternative forced-choice tests is mathematically equivalent to the area under the ROC curve from old/new tests with confidence ratings (Green and Moses, 1966; Green and Swets, 1966; Swets and Pickett, 1982; Stanislaw and Todorov, 1999). Swets and Pickett (1982) and Stanislaw and Todorov (1999) advocated for calculating the area under the ROC curve using maximum-likelihood estimation to fit the z-transformed ROC curve—a measure referred to as A_z—which does not assume equal variances of the target and distractor distributions. We reanalyzed published data from our laboratory (Experiment 4

in Stark et al., 2015) to examine whether there were differences in performance between the forced-choice test format and the old/new test format with confidence ratings. Specifically, if the forced-choice test format allowed participants to rely on familiarity to a greater extent than the old/new test format (Holdstock et al., 2002; Norman and O'Reilly, 2003; Migo et al., 2009, 2014), then we should observe significantly better performance on the forced choice test format than A_z from the old/new test format. Conversely, if the two test formats rely on similar mnemonic representations, then we should not observe a difference between the two test formats.

The difference between A_z calculated from the old/new test format for targets vs unrelated foils and the FC (A vs X) test format failed to reach significance. Similarly, the difference between A_z caluclated from the old/new test format for targets vs similar lures and FCC (A vs A') test format failed to reach significance. Moreover, we observed significantly worse performance on the FCNC (A vs B') test format than A_z from the old/new test format with targets vs similar lures. Taken together, these results suggest that the forced-choice format does not improve the discrimination between targets and unrelated foils nor the discrimination between targets and similar lures. We note that there were minor differences between the stimulus sets used in these experiments, however the results are at least consistent with the notion that forced-choice formats rely on the same mnemonic representations and do not receive familiarity-related enhancements in performance (cf. Khoe et al., 2000; Bayley et al., 2008; but see: Jeneson et al., 2010). Previous studies have shown similar performance on the FC (A vs X) test format and the old/new test format with targets and unrelated foils (Green and Moses, 1966; Khoe et al., 2000; Smith and Duncan, 2004). Bayley et al. (2008) showed that performance was similar on the FCC (A vs A') test format and the old/new test format with targets and similar lures in patients with selective hippocampal damage and in healthy control participants (but see: Jeneson et al., 2010). Future studies could assess the relationship between performance on the old/new test format with targets and similar lures (e.g., using A_z) and the FCC (A vs A') test format in more detail.

The results from the area under the ROC curve analysis can be brought to bear on our findings in healthy older adults. Specifically, it appears that the old/new discrimination between targets and similar lures most closely resembles the FCC (A vs A') test format. Previous studies from our laboratory and others have revealed an age-related impairment in the discrimination between targets and similar lures across a variety of test formats, including old/similar/new (Tonor et al., 2009; Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015), old/new (Stark et al., 2015), and old/new with confidence ratings (Stark et al., 2015). Moreover, these effect maintained across a variety of encoding conditions—e.g., incidental encoding, intentional encoding, continuous recognition (Stark et al., 2015). Altogether, there is an unequivocal age-related impairment in the ability to discriminate between targets and similar lures, which we argue is caused by a mnemonic rather than a decision-based difference between younger and healthy older adults. Indeed, the FCC (A vs A') test format eliminates any possible shifts in decision criterion across groups, similar to our previous report using ROC curves from the old/new test format with confidence ratings (Stark et al., 2015), thus obviating the concerns raised by Loiotile and Courtney (2015).

4.4.4 Application of global matching models to interpret other studies that used mnemonic similarity tasks

We also investigated whether global matching models could account for the results of other studies that have used mnemonic similarity tasks. For example, Reagh and Yassa (2014) used a variant of the Mnemonic Similarity Task to investigate the effect of stimulus repetition on memory for images of objects. They reported that stimulus repetition—three presentations compared to one presentation—improved discrimination between targets and unrelated foils and increased the false alarm rate to similar lures (using an old/new test format). They concluded that repetition improves generalization while impairing mnemonic discrimination. Moreover, they suggested that stimulus repetition can induce competition between memory

traces which would cause a loss of details from memory. Subsequently, Loiotile and Courtney (2015) used signal detection theory to show that while repetition increased the false alarm rate to similar lures, it also enhanced discrimination between targets and similar lures (as measured by d_a). They also showed that repetition improved performance on the FCC (A vs A') test format. The results from these studies were initially puzzling, and we were curious whether they could be accounted for within a global matching framework. To test this possibility, we modeled their tasks using MINERVA 2. Interestingly, our simulations have been able to account for the data from both studies. Specifically, MINERVA 2 predicts that repetition will cause: 1) better discrimination between targets and unrelated foils (as measured by an ROC analysis), 2) an increased false alarm rate to similar lures (cf. Hintzman, 1988, 2001; Hintzman et al., 1992), 3) better discrimination between targets and similar lures (as measured by an ROC analysis), 4) improved FCC (A vs A') test format performance.

The key insight from MINERVA 2 is that stimulus repetition increases the global match of similar lure items by increasing the number of traces that match the similar lure—i.e., three traces of A will generate a larger global match in response to A' than only a single trace. As a corollary, MINERVA 2 predicts that encoding the same exact details of an item three times would also increase the false alarm rate to a similar lure, suggesting that an increased false alarm rate to similar lures does not necessarily indicate a loss of details from memory. Furthermore, while repetition increases the global match of the similar lure distribution, it also decreases the overlap between the target and similar lure distributions. Therefore, the model predicts that comparisons between the target distribution and the similar lure distribution will be more discriminable for items that are presented three times than items that are presented one time (i.e., based on an ROC analysis or performance on the FCC (A vs A') test format). Altogether, the findings from our simulations highlight the notion that formal models can be used to constrain the interpretation of behavioral results. Thus, although global matching models have been challenged by a number of findings (for review see: Clark and Gronlund, 1996), we believe that they provide useful tools for interpreting

the results of studies that manipulate stimulus similarity.

Conclusion

Previous research has shown that there are clear age-related impairments on tasks that tax recollection and associative memory with a more mild impairment on tests of simpler item recognition memory (Craik and McDowd, 1987; Spencer and Raz, 1995; Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2004; Old and Naveh-Benjamin, 2008a,b; Danckert and Craik, 2013). Other studies have shown that healthy older adults reliably exhibit an impairment on item recognition memory tests that require discriminating between targets and similar lures (Tonor et al., 2009; Yassa et al., 2011b,a; Stark et al., 2013, 2015; Bennett et al., 2015). Our results suggest that healthy older adults are similarly impaired on the forced-choice discrimination between an object and its similar lure. Taken together, there is clear evidence that memory tests that require a high degree of fidelity are impaired in healthy older adults. Our modeling results suggest that aging causes an impaired ability to encode stimulus features, which causes fewer details to be encoded on each trial. These results provide a potential mechanistic interpretation of previous results that does not emphasize differences in cognitive processes but instead emphasizes differences in the mnemonic resolution required to solve the task (cf. Cowell et al., 2010a).

The present empirical and model-based results raise a number of interesting questions for future studies. First, future studies could use more stimuli in each forced-choice test format. For example, other reports have used four-alternative forced choice test formats (Migo et al., 2009, 2014). The four-alternative format provides a greater dynamic range than the two-alternative format used here. Although performance on the FCNC (A vs B') test format was above chance in both younger and older adults, it is possible that the four-alternative format would be more sensitive to detecting an age-related impairment. Similarly, a four-

alternative FC (A vs X) test format might be more senstive to detecting age-related changes in performance. Second, future studies could collect data from both the old/new test format (with confidence ratings) with targets and similar lures and the two-alternative FCC (A vs A') test format to investigate whether there is a significant relationship between A_z and proportion correct (across a pair of distinct but similarity-matched stimulus sets). Finally, future modeling work could be used to discover conditions in which global matching models (e.g., Hintzman, 1988; Murdock, 1995) and the dual-process complementary learning systems model (Norman and O'Reilly, 2003; Norman, 2010) generate disparate predictions, thus generating new ideas for behavioral tasks that directly test the predictions of the models.

Chapter 5

Conclusions

The ability to remember the events of our lives critically relies upon formation of associations among the "what," "where," and "when" aspects of the event. Following the discovery that damage to the hippocampus and adjacent medial temporal lobe cortical areas caused amnesia in human patients (Scoville and Milner, 1957), decades of research have focused on elucidating the mnemonic role of the medial temporal lobe (MTL). The results of these studies suggested that the MTL is necessary for declarative memory, including memory for facts and events, but that it plays little role in nondeclarative memory (e.g., Cohen and Squire, 1980; Squire, 1992; Squire and Zola, 1996; Eichenbaum, 2000). Many studies have investigated whether there are functional differences between the hippocampus and the adjacent MTL cortical regions.

A prominent theory suggests that the hippocampus is involved in recollection and the formation of associative memories (e.g., "what-where", "what-where-when") while the MTL cortical regions, and perirhinal cortex (PRC) in particular, are involved in familiarity (e.g., "what"; Brown and Aggleton, 2001; Yonelinas, 2002; Diana et al., 2007; Eichenbaum et al., 2007; Ranganath, 2010; Yonelinas et al., 2010). However, other theories suggest that the hip-

pocampus and MTL cortical regions are broadly involved in declarative memory, including recollection, familiarity, and associative and item-based memory (e.g., Squire et al., 2007; Wixted et al., 2010; Wixted and Squire, 2011). While there is disagreement regarding the localization of recollection (or associative memory) and familiarity (or item-based memory) to the hippocampus and PRC, respectively, proponents of both accounts have hypothesized that subregions of the MTL would differ in terms of the information that they represent (Eichenbaum et al., 2007; Diana et al., 2007; Wixted and Squire, 2011). Specifically, based on differences in anatomical connectivity between regions of the MTL, theoretical accounts have suggested that parahippocampal cortex (PHC; called "postrhinal cortex" in rodents) is involved in memory for contexts and spatial information while PRC is involved in memory for objects.

Recent theories suggest that the MTL-centered view of declarative memory is too limited and that we should expand our investigations to consider regions outside of the MTL. For example, many theories note a role for retrosplenial cortex (RSC) in the representation of contextual and spatial information, similar to PHC (Vann et al., 2009; Aggleton, 2010; Ranganath and Ritchey, 2012; Bucci and Robinson, 2014; Ritchey et al., 2015). Further, damage to RSC is known to cause "retrosplenial amnesia" (Valenstein et al., 1987; Aggleton, 2010) and a recent theory suggests that RSC is important for the formation of stimulus-stimulus associations (Bucci and Robinson, 2014), a function traditionally ascribed to the hippocampus. Thus, a more complete understanding of the involvement of structures beyond the MTL, especially RSC, is of keen interest. We were motivated by these theories; accordingly, four of the experiments in this dissertation were centered around investigating the representation of information in the MTL and retrosplenial cortex.

Previous research has suggested that healthy aging causes an impaired ability to form stimulus-stimulus associations (e.g., "what-where") with a relative sparing of performance of simpler item recognition memory (e.g., "what"; Craik and McDowd, 1987; Spencer and

Raz, 1995; Naveh-Benjamin, 2000; Naveh-Benjamin et al., 2004; Old and Naveh-Benjamin, 2008a,b; Danckert and Craik, 2013). Our laboratory has previously shown that healthy older adults are impaired in their ability to discriminate between previously viewed images of objects (i.e., targets) and similar lures—a task referred to as the Mnemonic Similarity Task (Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015). Previous studies have shown that there can be test-format effects on item recognition memory performance. In particular, the forced-choice test format (e.g., "Did you see object A or object B during the encoding phase?") can sometimes yield different results than tests that require old/new judgments of individual stimuli (e.g., "Did you see this object during the encoding phase?"). The forced-choice format relies on different assumptions (e.g., decision criteria) than the old/new test format; hence, converging evidence from these approaches would bolster the conclusion that healthy aging is accompanied by impaired mnemonic discrimination performance. Additionally, we used a class of models from mathematical psychology, global matching models, to investigate whether healthy aging could be modeled as an impaired ability to encode stimulus features. I will discuss these issues in the following sections.

5.1 Investigation of representations in PHC and RSC

In four functional magnetic resonance imaging (fMRI) experiments, we tested the hypothesis that PHC and RSC are preferentially involved in representing scenes and contextual information. We analyzed the data in all four experiments using multivariate pattern analysis (MVPA). As discussed in Chapter 1, MVPA relies on a different set of assumptions than activation analysis. Moreover, MVPA allowed us to investigate the representation of information in the MTL and RSC. In Chapter 2, we investigated the representation of images of faces and images of objects (Experiment 1) and of images of faces and scenes (Experiment 2). We tested the hypothesis that PHC and RSC are preferentially involved in the representation of

scenes, and we investigated whether they share information about scenes on a trial-by-trial basis. In Chapter 3, we aimed to build upon our findings in Chapter 2. Specifically, in Experiment 3, we investigated the representation of individual items and contexts. Finally, in Experiment 4, we investigated whether the representation of contexts was stable across different versions of the same context.

5.1.1 Investigation of scene representation in parahippocampal cortex and retrosplenial cortex

In Chapter 2, we discussed the results of two fMRI experiments. In Experiment 1, participants viewed images of faces and objects. In Experiment 2, participants viewed images of faces and scenes. We used a slow-event related design, which allowed us to analyze patterns of activity at the level of individual trials. We found that patterns of activity in PHC and RSC could be used to classify trials on which participants viewed images of faces and images of scenes (Experiment 1) as well as images of faces and images of scenes (Experiment 2). Classification accuracy was significantly better for the experiment with faces and scenes than the experiment with faces and objects, supporting the hypothesis that these regions are involved in scene processing. We next investigated whether these regions carried similar information about scenes on a trial-by-trial basis, which would be consistent with the notion that these regions share information about scenes.

Previous studies have investigated the similarity of the time-course of *activation* (and deactivation) of brain regions. These methods are largely referred to as functional connectivity methods because they attempt to investigate the functional coupling between given brain regions (based on co-activation; Friston, 1994). Coutanche and Thompson-Schill (2013) developed a technique, called informational connectivity, which investigates the similarity of the time-course of *information* contained within multiple brain regions. We used a trial-by-

trial variant of informational connectivity, which we refer to as informational correlativity. In short, our method calculated how "face-like" or how "object-like" (Experiment 1) and how "face-like" or how "scene-like" (Experiment 2) the patterns of activity were for each trial—a measure referred to as multivariate pattern discriminability. We then calculated the similarity of the trial-by-trial multivariate pattern discriminability between PHC and RSC. We hypothesized that if PHC and RSC share information about scenes, then we should observe stronger informational correlativity in our experiment with faces and scenes than in our experiment with faces and objects.

Consistent with our hypothesis, we observed significant informational correlativity between PHC and RSC in our experiment with faces and scenes but not in our experiment with faces and objects. Moreover, we observed significantly stronger informational correlativity between PHC and RSC in our experiment with faces and scenes than in our experiment with faces and objects. While our informational correlativity analysis was extremely limited in terms of the number of regions that we investigated, I think that this approach will be valuable for future studies. For example, Ritchey et al. (2014) applied a similar approach to investigate interactions in a much more global manner. Taken together, the results of our experiments support the hypothesis that PHC and RSC are preferentially involved in processing scenes, and that they share information in a stimulus-dependent manner (i.e., scenes). However, the finding that patterns of activity in both regions could be used to classify faces vs objects (i.e., non-scene images), suggests that a simple dissociation of scene or context processing is not sufficient to explain the information represented by PHC and RSC. In our next set of experiments, we aimed to further investigate the information processed by PHC and RSC.

5.1.2 What's in a context?

In Chapter 3, we discussed the results of Experiments 3 and 4, in which we investigated the representation of information during an associative memory task. The context-guided object association task has provided a useful framework to investigate the representations of events in the rodent (Rajji et al., 2006; Komorowski et al., 2009, 2013; Navawongse and Eichenbaum, 2013; Tort et al., 2013; McKenzie et al., 2014; Farovik et al., 2015; Keene et al., 2016; for review see: McKenzie et al., 2015). In this task, animals learn item-reward contingencies that differ based on the context, which is operationally defined as visually and tactilely distinct chambers. The results of these studies have firmly established the role of the hippocampus and MTL cortical regions in the processing of event information. In addition, these studies have shown that the hippocampus is necessary for task performance and that it represents not only spatial information (i.e., context, location) but also information about non-spatial information (i.e., valence, item).

McKenzie et al. (2014) and Keene et al. (2016) used a variant of MVPA, called representational similarity analysis, to investigate event representation in the rodent MTL. Briefly, representational similarity analysis calculates the similarity of spatial patterns of activity in response to different stimuli (see Chapter 1). Their results revealed that the hippocampus and adjacent MTL cortical regions represent information in a hierachically-organized manner based on differences in the stimuli that comprise an event. Most important for the present discussion was the finding that events that take place in a different context were represented by very different patterns of activity in the hippocampus and adjacent MTL cortex. We created two human versions of the context-guided object association task for use in fMRI experiments. In Experiment 3, we used distinct contexts (time lapse videos) and objects (images). A previously unaddressed question regarding the context-guided object association task is the extent to which the neural representations in the MTL are influenced by the low-level differences between stimuli in the context-guided object association task.

We addressed this question in Experiment 4, using a combined approach of stimulus filtering and model testing to eliminate low-level differences between our "contexts" and "objects." Specifically, we investigated whether the representations of contexts and objects are stable across different versions of the same contexts and objects.

Given the reliable findings of scene processing in PHC and RSC in Experiments 1 and 2, we aimed to investigate whether we would observe similar effects of context in a human version of the context-guided object association task. Additionally, given the robust representation of context in the rodent hippocampus, we aimed to provide evidence for the representation of context in the human hippocampus (for difficulties in decoding information from the human hippocampus see: Chapter 2; Diana et al., 2008; LaRocque et al., 2013; Op de Beeck et al., 2013). We analyzed our data using representational similarity analysis, similar to the studies in the rodent (McKenzie et al., 2014; Keene et al., 2016). Our results provide evidence that is consistent with the rodent studies; however, our results appear to be influenced by low-level differences between the stimuli that comprise the events.

Investigation of the representation of distinct contexts and objects

The results of Experiment 3 (with distinct stimuli) largely supported previous findings in the rodent. Specifically, our results revealed a significant relationship to the context matrix in PHC and RSC. A searchlight analysis revealed a significant relationship to the context matrix within a cluster of voxels in the left posterior hippocampus. Notably, McKenzie et al. (2014) found significant context representation in the rodent dorsal hippocampus (the rodent homolog of human posterior hippocampus) during performance of the context-guided object association task. Additionally, we performed informational correlativity analysis to investigate whether the left posterior hippocampus, PHC, and RSC contained similar representations of individual contexts on a trial-by-trial basis. We observed significant informational correlativity between all three regions, supporting the notion that they are part of a network

of regions involved in processing contextual information. Moreover, these results extend our previous finding of stimulus-dependent informational correlativity between PHC and RSC to the representation of individual contexts.

We also found evidence for conjunctive item-in-context representations in RSC, and representations in RSC were related to associative memory performance. These findings support recent theories that note a role for RSC in the formation of stimulus-stimulus associations (Aggleton, 2010; Bucci and Robinson, 2014). We also observed a clear relationship to the context matrix in parahippocampal place area and retrosplenial complex, which have both been extensively studied for their role in scene and context processing (e.g., Epstein and Kanwisher, 1998; Epstein et al., 2007; Julian et al., 2012; Vass and Epstein, 2013, 2016; Marchette et al., 2014, 2015). These regions are in close anatomical proximity to parahippocampal cortex and RSC but they are largely non-overlapping (see Fig. 3.5A). We also observed evidence for conjunctive item-in-context representation in parahippocampal place area and retrosplenial complex. Taken together, although our results were influenced by low-level visual features, these results suggest that RSC, parahippocampal place area, and retrosplenial complex represent information in addition to scenes and contexts. Notably, other studies have provided evidence for conjunctive representations in retrosplenial complex (e.g., Vass and Epstein, 2013; Marchette et al., 2014, 2015).

We observed clear evidence for a relationship between the V1 similarity matrix and the context matrix. Moreover, the V1 similarity matrix was related to the item-in-context matrix and the item-in-order-in-context matrix. In hindsight, this result is not surprising given the visual nature of the task. Interestingly, however, we observed a relationship between representations in V1 and associative memory performance. These findings highlight the difficulty of dissociating mnemonic from processing roles in experiments that use distinct contexts and objects. In the rodent version of the task, the contexts were comprised of chambers that differed in terms of visual and tactile cues and the objects differed in terms of visual, tactile,

and olfactory cues. Thus, it is possible that the results of these studies were also influenced by differences between the stimuli that comprise an event. However, other studies in the rodent have shown that more psychological versions of context can modulate place cell firing in the hippocampus—e.g., behavioral conditions (Wood et al., 2000; Smith and Mizumori, 2006b), expectations (Skaggs and McNaughton, 1998), temporal context (Allen et al., 2016). These results suggest that psychological context can dramatically influence representations in the rodent hippocampus even within the same physical environment (Smith and Mizumori, 2006a; Smith and Bulkin, 2014). We suggest, however, that future experiments should investigate the degree to which low-level features influence representations in the MTL. For example, studies that have recorded simultaneously from V1 and the hippocampus have observed "place cell" activity in V1 that has been shown to precede place cell activity in the hippocampus (Ji and Wilson, 2007; Haggerty and Ji, 2015). Accordingly, simultaneous recordings from multiple brain regions can elucidate the sensory and mnemonic contribution of regions of the MTL.

Given that whole-brain coverage is afforded by fMRI, human fMRI studies will be useful for testing the influence of low-level stimulus features on representations throughout the brain. However, the fact that we observed a relationship between representations in V1 and behavior highlights the importance of cross-species approaches. For example, optogenetic techniques in mice allow researchers to tag cells that are active during encoding and then to reactivate those cells under experimental control. Cowansage et al. (2014) tagged cells that were active in RSC during contextual fear conditioning. Similar to numerous reports, they showed that inactivation of the hippocampus reduced freezing behavior upon re-exposure to the context. Interestingly, optogenetic reactivation of cells in RSC during hippocampal inactivation was sufficient to elicit freezing behavior in a neutral context, suggesting that one role of the hippocampus in memory expression is to reactivate the relevant cortical patterns of activity (Teyler and DiScenna, 1986; Teyler and Rudy, 2007) in regions such as RSC. Thus, research in animal participants can provide more causal evidence for brain regions in

memory performance than the correlation approaches afforded by human neuroimaging.

The results of Experiment 3 should guide the experimental designs of future human neuroimaging studies. For example, our results highlight the necessity of investigating whether there are low-level visual confounds in the stimulus set before running an experiment that investigates the representation of context. In Experiment 4, we provided one approach for mitigating the influence of low-level visual features. Specifically, we used stimulus filtering (histogram matching) and model testing to select images that were devoid of a low-level visual confound. In Experiment 4, our contexts were multiple images of Saint Peter's Basilica and the U.S. Capitol Building and our objects were multiple images of car keys and house keys. We investigated whether the results from Experiment 3 would maintain after controlling for differences in the low-level visual features.

Investigation of invariant context and object representation

To verify that we adequately reduced the low-level differences between our categories of images, we first investigated representations in V1. Importantly, we saw no sign of a relationship between the similarity matrix in V1 and the context matrix and the object matrix. We used a computational model of V1 to show that there was a relationship between the similarity matrix from the model of V1 and the empirical V1 similarity matrix. These results serve as a positive control, establishing the quality of the data. We next investigated representations within the hippocampus, PHC, RSC, parahippocampal place area, and retrosplenial complex—regions that showed a large effect of context in Experiment 3. Unfortunately, the relationship between the invariant context matrix failed to exhibit a significant relationship to the similarity matrices in these regions. Importantly, these regions exhibited greater blood-oxygen-level dependent activation for events compared to the perceptual baseline task, suggesting that our regions were responding to the task. Taken together, these findings suggest that our results from Experiment 3 were influenced by the low-level stimulus differences

among the different events. In contrast to our relatively null findings for the relationship to the context matrix, Marchette et al. (2015) reported invariant representations across indoor and outdoor images of the same landmark in both parahippocampal place area and retrosplenial complex. They also provided evidence that representations in parahippocampal place area might be influenced by knowledge of the landmarks. These results raise the interesting possibility that we would have observed context processing in our regions of interest if our participants had real-world experience with our contexts. Future studies can investigate this possibility in more detail.

We found a significant relationship between the similarity matrix in PRC and the invariant object matrix. These results suggest that PRC exhibits similar patterns of activity across different versions of the same objects. We suggest that an important next step is to provide evidence for a double dissociation between PRC and V1. We first showed that V1 was significantly more related to the V1 model than PRC. We next showed that representations in PRC were significantly more related to the object matrix than V1. These results support the notion that PRC contains higher-level representations than V1. Moreover, these results suggest that PRC contains invariant representations of subordinate categories (i.e., car keys and house keys), supporting previous reports that have suggested that PRC is involved in processing fine-grained semantic information (Clarke and Tyler, 2014; for review see: Clarke and Tyler, 2015). Clarke and Tyler (2014) used a partial correlation approach in their investigation of peririnal cortex, which can be used account for different aspects of stimulus features. These approaches provide complementary information, thus both approaches will be useful for future studies that investigate the representation of objects and contexts.

5.1.3   Moving beyond simple dissociations

We tested the hypothesis that subregions of the MTL would differ in terms of the information that they represent (e.g., Diana et al., 2007; Eichenbaum et al., 2007; Wixted and Squire, 2011; Ranganath and Ritchey, 2012; Ritchey et al., 2015). The results from our experiments support the basic distinctions set forth in these theories while also suggesting that there will not be simple dissociations among these regions. For example, we showed that PRC exhibits invariant representations of objects. These findings accord with the notion that PRC plays a role in memory for objects. Additionally, in PHC and RSC, classification accuracy for faces and scenes was significantly greater than for faces and objects. Moreover, PHC and RSC shared information in a stimulus-dependent manner (i.e., scenes). We also showed significant classification of faces vs objects in PHC and RSC, suggesting that PHC and RSC are not tuned solely for spatial, contextual, or scene processing.

Other recent studies have provided mixed evidence for dissociations among MTL cortical regions. For example, Keene et al. (2016) showed that medial entorhinal cortex, a region hypothesized to be involved in spatial processing, carried information about non-spatial aspects of the context-guided object association task (i.e., object, valence). Additionally, they showed that PRC and lateral entorhinal cortex, regions hypothesized to be involved in memory for objects, carried information about non-object aspects of the task (i.e., context, location). Taken together, Keene et al. (2016) suggested that strong versions of the binding of items and contexts model (Diana et al., 2007; Eichenbaum et al., 2007) are incomplete. I share a similar view, and I think that the field of memory research has been too consumed by quests for dissociations. I am hopeful that future research will shift toward more integrative analyses to understand how multiple interacting regions give rise to memory. Approaches that combine MVPA and computational model testing hold great promise for understanding how information is processed in different brain regions. However, the results from our context-guided object association experiments raise important questions about the types of

experiments that we should be conducting. In the next section, I will discuss other findings from the field of visual neuroscience that have revealed that scene processing regions process low-level visual information to a greater degree than previously thought.

Studies from the field of visual neuroscience have challenged the interpretation of scene processing in "scene" processing regions

While many studies have investigated the involvement of regions such as parahippocampal place area and retrosplenial complex in the processing of scenes, few fMRI studies have investigated whether low-level stimulus properties are related to differences in the responses of these regions. Recent studies have found evidence for retinotopic processing in parahippocampal place area (Arcaro et al., 2009; Silson et al., 2015). Specifically, these studies observed that parahippocampal place area preferentially responded to stimuli presented in the upper contralateral visual field. Additionally, there have been reports that low-level spatial properties of stimuli can preferentially activate scene processing regions. For example, Rajimehr et al. (2011) showed that parahippocampal place area preferentially responded to images with high spatial frequencies. Additionally, Nasr and Tootell (2012) observed greater parahippocampal place area activation in response to stimuli with greater energy at cardinal orientations (i.e., vertical or horizontal) relative to oblique orientations. Nasr et al. (2014) developed a technique to measure the rectilinearity of images (i.e., the energy at 90 degree angles). They showed that scene stimuli from previously published studies had larger values of rectilinearity than stimuli from other categories. Moreover, they revealed greater parahippocampal place area activation for rectilinear stimuli compared to round stimuli. Finally, Lescroart et al. (2015) showed that a Fourier power model, which measures the spectral information of images, accounted for a large amount of variance in parahippocampal place area and retrosplenial cortex. Taken together, these findings highlight the challenges associated with concluding that regions play a role in *scene* or *object* processing. Moreover, these

results suggest that "high-level" cortical regions process low-level visual features.

In contrast to our approach for mitigating low-level differences between stimuli (Experiment 4), Lescroart et al. (2015) advocated for the use of very large stimulus sets. Moreover, they used encoding models to investigate the information that is represented in different voxels of the brain. Encoding models are relatively new techniques, but they provide a framework for linking information about stimulus features with the dynamics of activity in a given voxel (for review see: Naselaris et al., 2011). Specifically, estimates are generated for the features of each image in the stimulus set. Regression analysis is then used on a voxel-by-voxel basis to assign weights to the features of the encoding model (similar to activation analysis). Importantly, these techniques test the reliability of the encoding model using a left-out data set. These methods allow the investigation of the amount of variance accounted for by the model, thus providing information about the stimulus features that drive activity in a given voxel. The difficulty with applying encoding models to regions of the MTL, however, is that the features that will be represented in these regions are not well established, thus making it difficult to construct an encoding model. In contrast, the information processed in regions such as V1 are well established. However, as more data and theories accumulate, encoding models should be extended to investigate representations in the MTL. In the meantime, these findings highlight the notion that human fMRI studies that attempt to investigate category representation (including scenes, objects, and contexts) cannot ignore differences in low-level visual features between stimuli. A more holistic approach that incorporates the findings and methods from the field of visual neuroscience can help elucidate the information that is processed by regions of the MTL.

5.1.4 Future directions

In addition to the future directions discussed so far in this chapter, we listed several specific future directions for studies that aim to investigate the representation of context (see Chapter 3). I will reiterate those ideas here. First, rather than defining contexts and objects in an unconstrained manner, differences between contexts and objects could be manipulated to the same degree and within the same modality. One caveat, however, is that experiments that investigate differences in representations between two static stimuli will be virtually guaranteed to show differences in representations in the relevant sensory areas. However, changing the context and items to a similar degree would allow testing whether contexts or objects play a greater role in the representation of information in the MTL. It should be noted, however, that the objects themselves represent a context—i.e., the *context* of the object changes the predicted outcome of a behavioral response. Second, imaging/physiology data could be collected before and after training. Such studies should include control stimuli that consist of identically structured events but without a learned behavioral response or outcome. Third, the reverse type of experiment could also be used. Specifically, studies could investigate whether representations of distinct contexts become more similar to each other as a result of similar associations across the contexts; however, careful control will be required to show that there are not behavioral differences that emerge as a result of learning. Fourth, experiments could use the same physical stimuli with a different pre-event cue (cf. Skaggs and McNaughton, 1998; Wood et al., 2000; Smith and Mizumori, 2006a,b; Hsieh et al., 2014; Allen et al., 2016). Fifth, future studies could further investigate invariant representation, similar to our approach in Experiment 4. Finally, future studies could investigate the role of memory in invariant context and object representation.

5.2 Investigation of the effects of test format on performance of the Mnemonic Similarity Task

In Chapter 4, we investigated the effect of test format on performance of the Mnemonic Similarity Task in a group of younger adults and healthy older adults. Previous studies from our laboratory have revealed an age-related impairment in the ability to discriminate between targets and similar lures on both the old/similar/new test format and the old/new test format (Yassa et al., 2011a,b; Stark et al., 2013, 2015; Bennett et al., 2015). We tested the hypothesis that this effect would extend to a forced-choice variant of the Mnemonic Similarity Task. We used three different forced-choice test formats. In the first test format, participants were shown one target item and one unrelated foil item (A vs X). In the second test format, participants were shown one target item and a similar lure to that target item (A vs A'). In the third test format, participants were shown one target item and a similar lure item of a different target item (A vs B').

5.2.1 Forced-choice and old/new test formats reveal a stable age-related impairment of performance

In both younger and older adults, we observed an effect of test format. Specifically, performance was best on the A vs X test format, followed by the A vs A' test format, followed by the A vs B' test format. Moreover healthy older adults exhibited impaired performance on the A vs A' test format relative to younger adults. These results suggest that the previous reports of an age-related impairment in the ability to discriminate between targets and similar lures on the old/new and the old/similar/new test formats extends to the A vs A' test format. These test formats rely on a different set of assumptions; therefore, the consistent finding of an age-related impairment in the ability to discriminate between targets and sim-

ilar lures suggests a memory-based rather than a decision-based difference between the two groups, mitigating concerns raised by Loiotile and Courtney (2015) about age-related effects on performance of the Mnemonic Similarity Task.

5.2.2 Application of global matching models to the Mnemonic Similarity Task

We showed that a class of models from mathematical psychology, global matching models, provided a good qualitative fit to our empirical results in younger and older adults. The models predicted a similar effect of test format in both age groups. Performance on the A vs X test format was predicted to be the easiest due to the very low similarity between a target item and an unrelated foil. Performance on the A vs A' test format was predicted to be better than performance on the A vs B' test format. Previous studies have been able to account for this finding using global matching models (Hintzman, 1988, 2001; for review see: Clark and Gronlund, 1996). Specifically, these studies suggested that the shared variance between targets and lures on the A vs A' test format but not the A vs B' test format caused better performance on the A vs A' test format. We further investigated this effect by manipulating encoding variability.

In standard versions of MINERVA 2, each feature is encoded with a certain probability, however, there is trial-by-trial variability in the number of features that are encoded. We found that the removal of encoding variability reduced the A vs A' test format advantage. These results suggest that there are certain trials in which a participant happens to encode more features than others. Accordingly, there are trials on which a participant encodes more features of the original target of the noncorresponding lure (i.e., B) than of the target item (i.e., A). Because the similar lures are correlated with the target items, this results in certain trials where the global match of the B' item exceeds that of the A item. When the number

163

of encoded features is held constant, the model predicts a smaller difference in performance between the two tasks.

In our initial simulations we found parameters that fit the empirical data in younger adults. We next showed that reducing the probability of encoding stimulus features resulted in a similar pattern of results as we observed in older adults. Specifically, the models predicted that the largest change in performance would be on the A vs A' test format. Interestingly, however, the model predicted that there would also be decreases in performance on the A vs X test format and the A vs B' test format. Future studies could use more distractor images in each test format. For example, previous studies have used a four-alternative forced-choice test format, which has a larger dynamic range than our two-alternative forced-choice test formats. The use of the four-alternative forced-choice test format might be more sensitive for detecting age-related differences on both the A vs X test format and the A vs B' test format. Future studies could also further investigate the relationship between performance on old/new and forced-choice test formats. Finally, future studies could further explore conditions in which global matching models and the dual-process complementary learning systems model (Norman and O'Reilly, 2003) make disparate predictions about behavior.

In addition to their ability to account for our data, we also showed that global matching models provided a possible mechanistic account of other studies that have used the Mnemonic Similarity Task (Reagh and Yassa, 2014; Loiotile and Courtney, 2015; see Chapter 4). In conclusion, the global matching framework provided a reasonable account of our empirical findings. Importantly, our use of the global matching models emphasized differences in the mnemonic resolution required to solve the different test formats rather than differences in cognitive processes (cf. Cowell et al., 2010a,b). In the final section, I will briefly discuss the potential for combining functional analysis and cognitive modeling.

5.2.3   A combined functional and model-based approach

Recent studies have combined cognitive modeling and neuroimaging. For example, Davis et al. (2014) used a global matching model to analyze fMRI data that were collected as participants performed a recognition memory task. Their modeling approach was similar to our approach in Chapter 4. However, instead of using randomly generated vectors to represent items, they used patterns of activity from the MTL. Interestingly, they showed that the global match between the pattern of activity of a given target and the patterns of activity of all of the encoding trials was related to recognition memory performance. These results directly support the global matching framework by suggesting that recognition memory decisions are influenced by the global match between a target item and the items in memory. Other studies have used fMRI to detect latent cognitive processes. For example, Gershman et al. (2013) used MVPA to support the theory that the reinstatement of a previous context during new learning was related to source memory errors. Similarly, Polyn et al. (2005) demonstrated that patterns of activity during the recollection of an item were similar to patterns of activity during the initial encoding phase, which supports the hypothesis that recollection reactivates the context in which an item was studied. These types of experiments are interesting because they use fMRI as a tool to examine cognitive states rather than attempts to map where in the brain certain cognitive functions take place. Taken together, the combination of modeling (both computational and cognitive) and MVPA will be a fruitful framework for future studies.

References

Aggleton, J. P. (2010). Understanding retrosplenial amnesia: insights from animal studies. *Neuropsychologia*, 48:2328–2338.

Aggleton, J. P. and Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences*, 22:425–489.

Aguirre, G. K. (2007). Continuous carry-over designs for fmri. *NeuroImage*, 35(4):1480–1494.

Aguirre, G. K., Mattar, M. G., and Magis-Weinberg, L. (2011). de Bruijn cycles for neural decoding. *NeuroImage*, 56(3):1293–1300.

Alexander, A. S. and Nitz, D. A. (2015). Retrosplenial cortex maps the conjunction of internal and external spaces. *Nature Neuroscience*, 18(8):1143–1151.

Allen, T. A., Salz, D. M., McKenzie, S., and Fortin, N. J. (2016). Nonspatial sequence coding in CA1 neurons. *The Journal of Neuroscience*, 36(5):1547–1563.

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

Arcaro, M. J., McMains, S. A., Singer, B. D., and Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *The Journal of Neuroscience*, 29(34):10638–10652.

Auger, S. D. and Maguire, E. A. (2013). Assessing the mechanism of response in the retrosplenial cortex of good and poor navigators. *Cortex*, 49(10):2904–2913.

Auger, S. D., Zeidman, P., and Maguire, E. A. (2015). A central role for retrosplenial cortex in de novo enviromental learning. *eLife*, 4:e09031.

Avants, B., Duda, J. T., Kim, J., Zhang, H. Y., Pluta, J., Gee, J. C., and Whyte, J. (2008). Multivariate analysis of structural and diffusion imaging in traumatic brain injury. *Academic Radiology*, 15:1360–1375.

Azab, M., Stark, S. M., and Stark, C. E. L. (2014). Contributions of human hippocampal subfields to spatial and temporal pattern separation. *Hippocampus*, 24(3):293–302.

Bakker, A., Kirwan, C. B., Miller, M. I., and Stark, C. E. L. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*, 319(5870):1640–1642.

Bar, M. and Aminoff, E. (2003). Cortical analysis of visual context. *Neuron*, 38(2):347–358.

Barnes, C. A., McNaughton, B. L., Misumori, S. J., Leonard, B. W., and Lin, L. H. (1990). Comparison of spatial and temporal characteristics of neuronal activity in sequential stages of hippocampal processing. *Progress in Brain Research*, 83:287–300.

Barnes, J. M. and Underwood, B. J. (1959). Fate of first-list associations in transfer theory. *Journal of Experimental Psychology*, 58:97–105.

Bayley, P. J., Wixted, J. T., Hopkins, R. O., and Squire, L. R. (2008). Yes/no recognition, forced-choice recognition, and the human hippocampus. *The Journal of Cognitive Neuroscience*, 20:505–512.

Bennett, I. J., Huffman, D. J., and Stark, C. E. L. (2015). Limbic tract integrity contributes to pattern separation performance across the lifespan. *Cerebral Cortex*, 25(5):2988–2999.

Bonnici, H. M., Kumaran, D., Chadwick, M. J., Weiskopf, N., Hassabis, D., and Maguire, E. A. (2012). Decoding representations of scenes in the medial temporal lobes. *Hippocampus*, 22:1143–1153.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10:443–446.

Brown, M. W. and Aggleton, J. P. (2001). Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, 2:51–61.

Brown, M. W., Warburton, E. C., and Aggleton, J. P. (2010). Recognition memory: material, processes, and substrates. *Hippocampus*, 20:1228–1244.

Bucci, D. J. and Robinson, S. (2014). Toward a conceptualization of retrohippocampal contributions to learning and memory. *Neurobiology of Learning and Memory*, 116:197–207.

Buckner, R. L., Snyder, A. Z., Shannon, B. J., LaRossa, G., Sachs, R., Fotenos, A. F., Sheline, Y. I., Klunk, W. E., Mathis, C. A., Morris, J. C., and Mintun, M. A. (2005). Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *The Journal of Neuroscience*, 25(34):7709–7717.

Buffalo, E. A., Bellgowan, P. S. F., and Martin, A. (2006). Distinct roles for medial temporal lobe structures in memory for objects and their locations. *Learning and Memory*, 13:638–643.

Bulkin, D. A., Law, M., and Smith, D. M. (2016). Placing memories in context: hippocampal representations promote retrieval of appropriate memories. *Hippocampus*, doi: 10.1002/hipo.22579.

Burke, S. N., Ryan, L., and Barnes, C. A. (2012). Characterizing cognitive aging of recognition memory and related processes in animal models and in humans. *Frontiers in Aging Neuroscience*, 4(15):1–13.

Burwell, R. D. (2000). The parahippocampal region: corticocortical connectivity. *Annals of the New York Academy of Sciences*, 911:25–42.

Chadwick, M. J., Bonnici, H. M., and Maguire, E. A. (2012). Decoding information in the human hippocampus: a user's guide. *Neuropsychologia*, 50:3107–3121.

Chadwick, M. J., Hassabis, D., and Maguire, E. A. (2011). Decoding overlapping memories in the medial temporal lobes using high-resolution fMRI. *Learning and Memory*, 18:742–746.

Chadwick, M. J., Hassabis, D., Weiskopf, N., and Maguire, E. A. (2010). Decoding individual memory traces in the human hippocampus. *Current Biology*, 20:544–547.

Chawla, M. K., Guzowski, J. F., Ramirez-Amaya, V., Lipa, P., Hoffman, K. L., Marriott, L. K., Worley, P. F., McNaughton, B. L., and Barnes, C. A. (2005). Sparse, environmentally selective expression of arc rna in the upper blade of the rodent fascia dentata by brief spatial experience. *Hippocampus*, 15:579–586.

Chen, L. L., Lin, L.-H., Green, E. J., A, B. C., and McNaughton, B. L. (1994). Head-direction cells in the rat posterior cortex. i. anatomical and behavioral modulation. *Experimental Brain Research*, 101:8–23.

Chen, Y., Namburi, P., Elliot, L. T., Heinzie, J., Soon, C. S., Chee, M. W., and Haynes, J. D. (2011). Cortical surface-based searchlight decoding. *NeuroImage*, 56(2):582–592.

Cho, J. and Sharp, P. E. (2001). Head direction, place, and movement correlates for cells in the rat retrosplenial cortex. *Behavioral Neuroscience*, 115(1):3–25.

Churchland, P. M. (1986). Some reductive strategies in cognitive neurobiology. *Mind*, 95:279–309.

Clark, S. E. and Gronlund, S. D. (1996). Global matching models of recognition memory: how the models fit the data. *Psychonomic Bulletin & Review*, 3(1):37–60.

Clarke, A. and Tyler, L. K. (2014). Object-specific semantic coding in human perirhinal cortex. *The Journal of Neuroscience*, 34(14):4766–4775.

Clarke, A. and Tyler, L. K. (2015). Understanding what we see: how we derive meaning from vision. *Trends in Cognitive Sciences*, 19(11):677–687.

Cohen, N. J., Ryan, J., Hunt, C., Romine, L., Wszalek, T., and Nash, C. (1999). Hippocampal system and declarative (relational) memory: summarizing the data from functional neuroimaging studies. *Hippocampus*, 9(1):83–98.

Cohen, N. J. and Squire, L. R. (1980). Preserved learning and retention of pattern analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science*, 210:207–209.

Coutanche, M. N. and Thompson-Schill, S. L. (2012). The advantage of brief fMRI acquisition runs for multi-voxel pattern detection across runs. *NeuroImage*, 61(4):1113–1119.

Coutanche, M. N. and Thompson-Schill, S. L. (2013). Informational connectivity: identifying synchronized discriminability of multi-voxel patterns across the brain. *Frontiers in Human Neuroscience*, 7(15):1–14.

Cowansage, K. K., Shuman, T., Dillingham, B. C., Chang, A., Colshani, P., and Mayford, M. (2014). Direct reactivation of a coherent neocortical memory of context. *Neuron*, 84(2):432–441.

Cowell, R. A., Bussey, T. J., and Saksida, L. M. (2010a). Components of recognition memory: dissociable cognitive processes or just differences in representational complexity? *Hippocampus*, 20(11):1245–1262.

Cowell, R. A., Bussey, T. J., and Saksida, L. M. (2010b). Functional dissociations within the ventral object processing pathway: cognitive modules or a hierarchical continuum? *Journal of Cognitive Neuroscience*, 22(11):2460–2479.

Cowell, R. A., Bussey, T. J., and Saksida, L. M. (2012). Empiricists are from Venus, modelers are from Mars: Reconciling experimental and computational approaches in cognitive neuroscience. *Neuroscience and Biobehavioral Reviews*, 36:2371–2379.

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computational Biomedical Research*, 29(3):162–173.

Craik, F. I. M. and McDowd, J. M. (1987). Age differences in recall and recogition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(3):474–479.

Crum, R. M., Anthony, J. C., Bassett, S. S., and Folstein, M. F. (1993). Population-based norms for the Mini-Mental State Examination by age and education level. *The Journal of the American Medical Association*, 269(18):2386–2391.

Cutzu, F. and Edelman, S. (1998). Representation of object similarity in human vision: psychophysics and a computational model. *Vision Research*, 38:2229–2257.

Danckert, S. L. and Craik, F. I. M. (2013). Does aging affect recall more than recognition memory? *Psychology and Aging*, 28(4):902–909.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, 16:693–700.

Davis, T. and Poldrack, R. A. (2013). Measuring neural representations with fMRI: practices and pitfalls. *Annals of the New York Academy of Sciences*, 1296:108–134.

Davis, T., Xue, G., Love, B. C., Preston, A. R., and Poldrack, R. A. (2014). Global neural similarity as a common basis for categorization and recognition memory. *The Journal of Neuroscience*, 34(22):7472–7484.

Desgranges, B., Baron, J. C., Lalevée, C., Giffard, B., Viader, F., de La Sayette, V., and Eustache, F. (2002). The neural substrates of episodic memory impairment in Alzheimer's disease as revealed by FDG-PET: relationship to the degree of deterioration. *Brain*, 125(5):1116–1124.

169

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., and Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31:968–980.

Destrieux, C., Fischl, B., Dale, A., and Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, 53(1):1–15.

Diana, R. A., Yonelinas, A. P., and Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in Cognitive Sciences*, 11(9):379–386.

Diana, R. A., Yonelinas, A. P., and Ranganath, C. (2008). High-resolution multi-voxel pattern analysis of category selectivity in the medial temporal lobe. *Hippocampus*, 18:536–541.

Duff, M. C., Warren, D. E., Gupta, R., Vidal, J. P. B., Tranel, D., and Cohen, N. J. (2012). Teasing apart the tangrams: testing hippocampal pattern separation with a collaborative referencing paradigm. *Hippocampus*, 22:1087–1091.

Edelman, S. (1998). Representation is representation of similarities. *Behavioral Brain Science*, 21(4):449–467.

Edelman, S. (2012). Renewing the respect for similarity. *Frontiers in Computational Neuroscience*, 6(45):1–19.

Edelman, S., Grill-Spector, K., Kushnir, T., and Malach, R. (1998). Toward direct visualization of the internal shape representation space by fMRI. *Psychobiology*, 26:309–321.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nature Reviews Neuroscience*, 1(1):41–50.

Eichenbaum, H. and Cohen, N. J. (2001). *From conditioning to conscious recollection: memory systems in the brain.* Oxford University Press.

Eichenbaum, H. and Cohen, N. J. (2014). Can we reconcile the declarative memory and spatial navigation views of the hippocampus? *Neuron*, 83(4):764–770.

Eichenbaum, H., Yonelinas, A. P., and Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, 30:123–152.

Ennaceur, A., Neave, N., and Aggleton, J. P. (1997). Spontaneous object recognition and object location memory in rats: the effects of lesions in the cingulate cortices, the medial prefrontal cortex, the cingulum bundle and the fornix. *Experimental Brain Research*, 113:509–519.

Epstein, R., Harris, A., Stanley, D., and Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, 23:115–125.

Epstein, R. and Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392:598–601.

Epstein, R. A., Parker, W. E., and Feiler, A. M. (2007). Where am I now? distinct roles for parahippocampal and retrosplenial cortices in place recognition. *The Journal of Neuroscience*, 27(23):6141–6149.

Ernst, M. D. (2004). Permutation methods: a basis for exact inference. *Statistical Science*, 19:676–685.

Etzel, J. A. (2015). MVPA permutation schemes: permutation testing for the group level. *Pattern Recognition in NeuroImaging (PRNI), 2015 International Workshop on*, pages 65–68.

Etzel, J. A., Gazzola, V., and Keysers, C. (2009). An introduction to anatomical ROI-based fMRI classification analysis. *Brain Research*, 1282:114–125.

Farovik, A., Place, R. J., McKenzie, S., Porter, B., Munro, C. E., and Eichenbaum, H. (2015). Orbitofrontal cortex encodes memories within value-based schemas and represents contexts that guide memory retrieval. *The Journal of Neuroscience*, 35(21):8333–8344.

Fennema-Notestine, C., Hagler, D. J. J., McEvoy, L. K., Wu, E. H., Karow, D. S., Dale, A. M., and Initiative, A. D. N. (2009). Structural MRI biomarkers for preclinical and mild Alzheimer's disease. *Human Brain Mapping*, 30(10):3238–3253.

French, R. M. (1991). Using semi-distributed representations to overcome catastrophic forgetting in connectionist networks. In *Proceedings of the 13th Annual Cognitive Science Society Conference*, pages 173–178. New Jersey: Lawrence Erlbaum.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human Brain Mapping*, 2:56–78.

Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. J., and Dolan, R. J. (1996). The trouble with cognitive subtraction. *NeuroImage*, 4:97–104.

Gener, T., Perez-Mendez, L., and Sanchez-Vives, M. V. (2013). Tactile modulation of hippocampal place fields. *Hippocampus*, 23(12):1453–1462.

Gershman, S. J., Schapiro, A. C., Hupbach, A., and Norman, K. A. (2013). Neural context reinstatement predicts memory misattribution. *The Journal of Neuroscience*, 33(20):8590–8595.

Gilbert, P. E., Kesner, R. P., and DeCoteau, W. E. (1998). Memory for spatial location: role of the hippocampus in mediating spatial pattern separation. *Journal of Neuroscience*, 18(2):804–810.

Gilbert, P. E., Kesner, R. P., and Lee, I. (2001). Dissociating hippocampal subregions: a double dissociation between dentate gyrus and CA1. *Hippocampus*, 11:626–636.

Goodale, M. A. and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, 15(1):20–25.

Green, D. M. and Moses, F. L. (1966). On the equivalence of two recognition memory measures of short-term memory. *Psychological Bulletin*, 66(3):228–234.

Green, D. M. and Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley & Sons.

Griffin, G., Holub, A., and Perona, P. (2007). Caltech-256 object category dataset. *Technical Report 7694, California Institute of Technology, 2007*.

Haggerty, D. C. and Ji, D. (2015). Activities of visual cortical and hippocampal neurons co-fluctuate in freely moving rats during spatial behavior. *eLife*, doi: 10.7554/eLife.08902.

Hanke, M. and Halchenko, Y. O. (2012). Neuroscience runs on GNU/Linux. *Neuroinformatics*, Jul 7:5–8.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., and S, P. (2009). PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, 7(1):37–53.

Hargreaves, E. L., Mattfeld, A. T., Stark, C. E. L., and Suzuki, W. A. (2012). Conserved fMRI and LFP signals during new associative learning in the huamn and macaque monkey medial temporal lobe. *Neuron*, 74:743–752.

Hassabis, D., Chu, C., Rees, G., Weiskopf, N., Molyneux, P. D., and Maguire, E. A. (2009). Decoding neuronal ensembles in the human hippocampus. *Current Biology*, 19:546–554.

Hasselmo, M. E., Schnell, E., and Barkai, E. (1995). Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in hippocampal region CA3. *Journal of Neuroscience*, 15(7):5249–5262.

Hasselmo, M. E. and Wyble, B. P. (1997). Simulation of the effects of scopolamine on free recall and recognition in a network model of the hippocampus. *Behavioral Brain Research*, 89:1–34.

Hasselmo, M. E., Wyble, B. P., and Wallenstein, G. V. (1996). Encoding and retrieval of episodic memories: role of cholinergic and gabaergic modulation in the hippocampus. *Hippocampus*, 6(6):693–708.

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: the early beginnings. *NeuroImage*, 62:852–855.

Haxby, J. V., Gobbini, M. I., Lurey, M. L., A, I., Schouten, J. L., and P, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cotex. *Science*, 293:2425–2430.

Haynes, J. D. and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7):523–534.

Hinds, O. P., Rajendran, N., Polimeni, J. R., Augustinack, J. C., Wiggins, G., Wald, L. L., Rosas, H. D., Potthast, A., Schwarts, E. L., and Fischl, B. (2008). Accurate prediction of V1 location from cortical folds in a surface coordinate system. *NeuroImage*, 39(4):1585–1599.

Hinton, G. F. (1981). Shape representation in parallel systems. In *Proceedings of the Seventh Annual International Joint Conference on Artificial Intelligence*, pages 1088–1096. Vancouver, BC, Canada.

Hintzman, D. L. (1984). MINERVA 2: a simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, 16(2):96–101.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95(4):528–551.

Hintzman, D. L. (2001). Similarity, global matching, and judgments of frequency. *Memory & Cognition*, 29(4):547–556.

Hintzman, D. L., Curran, T., and Oppy, B. (1992). Effects of similarity and repetition on memory: registration without learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(4):667–680.

Ho, A. S., Etsuro, H., Nguyen, P. H. T., Urakawa, S., Kondoh, T., Torii, K., Ono, T., and Nishijo, H. (2011). Hippocampal neuronal responses signaled licking of gustatory stimuli in different contexts. *Hippocampus*, 21(5):502–519.

Holdstock, J. S., Mayes, A. R., Gong, Q. Y., Roberts, N., and Kapur, N. (2005). Item recognition is less impaired than recall and associative recognition in patients with selective hippocampal damage. *Hippocampus*, 15(2):203–215.

Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., and Norman, K. A. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus*, 12:341–351.

Hsieh, L. T., Gruber, M. J., Jenkins, L. J., and Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron*, 81(5):1165–1178.

Huffman, D. J. and 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.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., Laakso, M. P., and Pitknen, A. (1998). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology*, 19:659–671.

Itskov, P. M., Vinnik, E., Honey, C., Schnupp, J., and Diamond, M. E. (2012). Sound sensitivity of neurons in rat hippocampus during performance of a sound-guided task. *Journal of Neurophysiology*, 107(7):1822–1834.

Jeneson, A., Kirwan, C. B., Hopkins, R. O., Wixted, J. T., and Squire, L. R. (2010). Recognition memory and the hippocampus: a test of the hippocampal contribution to recollection and familiarity. *Learning & Memory*, 17(1):63–70.

Ji, D. and Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience*, 10(1):100–107.

Julian, J. B., Fedorenko, E., Webster, J., and Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage*, 60(4):2357–2364.

Kahana, M. J. (2012). *Foundations of human memory*. Oxford University Press.

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11):4302–4311.

Kanwisher, N., Stanley, D., and Harris, A. (1999). The fusiform face area is selective for faces not animals. *Neuroreport*, 10(1):183–187.

Keene, C. S., Bladon, J., McKenzie, S., Liu, C. D., O'Keefe, J., and Eichenbaum, H. (2016). Complimentary functional organization of neuronal activity patterns in the perirhinal, lateral entorhinal, and medial entorhinal cortices. *The Journal of Neuroscience*, 36(13):3660–3675.

Khoe, W., Kroll, N. E. A., Yonelinas, A. P., Dobbins, I. G., and Knight, R. T. (2000). The contribution of recollection and familiarity to yes-no and forced-choice recognition tests in healthy subjects and amnesics. *Neuropsychologia*, 38(10):1333–1341.

Kirwan, C. B., Hartshorn, A., Stark, S. M., Goodrich-Hunsaker, N. J., Hopkins, R. O., and Stark, C. E. L. (2012). Pattern separation deficits following damage to the hippocampus. *Neuropsychologia*, 50(10):2408–2414.

Kirwan, C. B. and Stark, C. E. L. (2004). Medial temporal lobe activation during encoding and retrieval of novel face-name pairs. *Hippocampus*, 14:919–930.

Kirwan, C. B. and Stark, C. E. L. (2007). Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. *Learning and Memory*, 14:625–633.

Kirwan, C. B., Wixted, J. T., and Squire, L. R. (2008). Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. *Journal of Neuroscience*, 28(42):10541–10548.

Kivisaari, S. L., Tyler, L. K., Monsch, A. U., and Taylor, K. I. (2012). Medial perirhinal cortex disambiguates confusable objects. *Brain*, 135(Pt 12):3757–3769.

Knierim, J. J., Lee, I., and Hargreaves, E. L. (2006). Hippocampal place cells: parallel input streams, subregional processing, and implications for episodic memory. *Hippocampus*, 16:755–764.

Knierim, J. J., Neunuebel, J. P., and Deshmukh, S. S. (2013). Functional correlates of the lateral and medial entorhinal cortex: objects, path integration, and local-global reference frames. *Philosophical Transactions of the Royal Society B*, 369(1635):20130369.

Kolarik, B. S., Shahlaie, K., Hassan, A., Borders, A. A., Kaufman, K. C., Gurkoff, G., Yonelinas, A. P., and Ekstrom, A. D. (2016). Impairments in precision, rather than spatial strategy, characterize performance on the Virtual Water Maze: a case study. *Neuropsychologia*, 80:90–101.

Komorowski, R. W., Garcia, C. G., Wilson, A., Hattori, S., Howard, M. W., and Eichenbaum, H. (2013). Ventral hippocampal neurons are shaped by experience to represent behaviorally relevant contexts. *The Journal of Neuroscience*, 33(18):8079–8087.

Komorowski, R. W., Manns, J. R., and Eichenbaum, H. (2009). Robust conjunctive item-place coding by hippocampal neurons parallels what happens where. *The Journal of Neuroscience*, 29(31):9918–9928.

Kriegeskorte, N. (2009). Relating population-code representations between man, monkey, and computational models. *Frontiers in Neuroscience*, 3:363–373.

Kriegeskorte, N., Mur, M., and Bandettini, P. (2008a). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2(4).

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., and Bandettini, P. A. (2008b). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, 60(6):1126–1141.

Krzanowski, W. J. (2000). *Principles of multivariate analysis: a user's perspective*. Oxford University Press.

Kuhl, B. A., Rissman, J., and Wagner, A. D. (2011). Fidelity of neural reactivation reveals competition between memories. *Proceedings of the National Academy of the Sciences USA*, 108(14):5903–5908.

Kuhl, B. A., Rissman, J., and Wagner, A. D. (2012). Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia*, 50(4):458–469.

Kumaran, D. and Maguire, E. A. (2009). Novelty signals: a window into hippocampal information processing. *Trends in Cognitive Sciences*, 13(2):47–54.

Lacy, J. W., Yassa, M. A., Stark, S. M., and Stark, C. E. L. (2011). Distinct pattern separation related transfer functions in human CA1 and CA3/DG using high-resolution fMRI and mnemonic similarity. *Learning and Memory*, 18:15–18.

LaRocque, K. F., Smith, M. E., Carr, V. A., Witthoft, N., Grill-Spector, K., and 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.

Lavenex, P. and Amaral, D. G. (2000). Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus*, 10:420–430.

Law, J. R., Flanery, M. A., Wirth, S. W., Yanike, M., Smith, A. C., Frank, L. M., Suzuki, W. A., Brown, E. N., and Stark, C. E. L. (2005). Functional magnetic resonance imaging activity during the gradual acquisition and expression of paired-associate memory. *The Journal of Neuroscience*, 25(24):5720–5729.

Lazebnik, S., Schmid, C., and Ponce, J. (2006). Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. *2006 IEEE Computer Science Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, 2:2169–2178.

Lescroart, M. D., Stansbury, D. E., and Gallant, J. L. (2015). Fourier power, subjective distance, and object categories all provide plausible models of BOLD responses in scene-selective areas. *Frontiers in Computational Neuroscience*, 9:135.

Leutgeb, J. K., Leutgeb, S., Moser, M. B., and Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science*, 315(5814):961–966.

Leutgeb, J. K., Leutgeb, S., Treves, A., Meyer, R., Barnes, C. A., McNaughton, B. L., Moser, M. B., and Moser, E. I. (2005). Progressive transformation of hippocampal neuronal representations in "morphed" environments. *Neuron*, 48(2):345–358.

Leutgeb, S. and Leutgeb, J. K. (2007). Pattern separation, pattern completion, and new neuronal codes within a continuous CA3 map. *Learning and Memory*, 14(11):745–757.

Leutgeb, S., Leutgeb, J. K., Treves, A., Moser, M. B., and Moser, E. I. (2004). Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science*, 305:1295–1298.

Liang, J. C., Wagner, A. D., and Preston, A. R. (2013a). Content representation in the human medial temporal lobe. *Cerebral Cortex*, 23(1):80–96.

Liang, M., Mouraux, A., Hu, L., and Iannetti, G. D. (2013b). Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. *Nature Communications*, 4:1979.

Litman, L., Awipi, A., and Davachi, L. (2009). Category-specificity in the human medial temporal lobe cortex. *Hippocampus*, 19:308–319.

Loiotile, R. E. and Courtney, S. M. (2015). A signal detection theory analysis of behavioral pattern separation paradigms. *Learning & Memory*, 22(8):364–369.

Mandler, G. (1980). Recognizing: the judgment of previous occurence. *Psychological Review*, 87(3):252–271.

Marchette, S. A., Vass, L. K., Ryan, J., and Epstein, R. (2015). Outside looking in: landmark generalization in the human navigational system. *The Journal of Neuroscience*, 35(44):14896–14908.

Marchette, S. A., Vass, L. K., Ryan, J., and Epstein, R. A. (2014). Anchoring the neural compass: coding of local spatial reference frames in human medial parietal lobe. *Nature Neuroscience*, 17(11):1580–1606.

Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical Transactions of the Royal Society B*, 262(841):23–81.

Martin, A. (1999). Automatic activation of the medial temporal lobe during encoding: lateralized influences of meaning and novelty. *Hippocampus*, 9:62–70.

Mattfeld, A. T. and Stark, C. E. L. (2010). Striatal and medial temporal lobe functional interactions during visuomotor associative learning. *Cerebral Cortex*, 21(3):647–658.

Mattfeld, A. T. and Stark, C. E. L. (2011). Functional specialization within the striatum along both the dorsal/ventral and anterior/posterior axes during associative learning via reward and punishment. *Learning and Memory*, 18:703–711.

Mayes, A. R., Holdstock, J. S., Isaac, C. L., Hunkin, N. M., and Roberts, N. (2002). Relative sparing of item recognition memory in a patient with adult-onset damage limited to the hippocampus. *Hippocampus*, 12:325–340.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–457.

McCloskey, M. and Cohen, N. (1989). Catastrophic interference in connectionist networks: the sequential learning problem. In *The Psychology of Learning and Motivation*, volume 24, pages 109–164. NY: Academic Press.

McKenzie, S., Frank, A. J., Kinsky, N. R., Porter, B., and Rivière, P. D. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron*, 83(1):202–215.

McKenzie, S., Keene, C. S., Farovik, A., Bladon, J., Place, R., Komorowski, R., and Eichenbaum, H. (2015). Representation of memories in the cortical-hippocampal system: results from the application of population similarity analyses. *Neurobiology of Learning and Memory*, S1074-7427(15):00238–5.

McNaughton, B. L. and Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences*, 10(10):408–415.

Migo, E., Montaldi, D., Norman, K. A., Quamme, J., and Mayes, A. R. (2009). The contribution of familiarity to recognition memory is a function of test format when using similar foils. *The Quarterly Journal of Experimental Psychology*, 62:1198–1215.

Migo, E. M., Quamme, J. R., Holmes, S., Bendell, A., Norman, K. A., Mayes, A. R., and Montaldi, D. (2014). Individual differences in forced-choice recognition memory: partitioning contributions of recollection and familiarity. *The Quarterly Journal of Experimental Psychology*, 67(11):2189–2206.

Milner, B. (1962). Les troubles de la memoire accompagnant des lesions hippocampiques bilaterales. In Passouant, P., editor, *Physiologie de l'hippocampe*, pages 257–272. Paris: Centre National de la Recherche Scientifique.

Milner, B., Corkin, S., and Teuber, H. L. (1968). Further analysis of the hippocampal amnesiac syndrome: 14-year follow-up study of H.M. *Neuropsychologia*, 6:215–234.

Minoshima, S., Giordani, B., Berent, S., Frey, K. A., Foster, N. L., and Kuhl, D. E. (1997). Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Annals of Neurology*, 42(1):85–94.

Mishkin, M., Suzuki, W. A., Gadian, D. G., and Vargha-Khadem, F. (1997). Hierarchical organization of cognitive memory. *Philisophical Transactions of the Royal Society B*, 352:1461–1467.

Morgan, L. K., MacEvoy, S. P., Aguirre, G. K., and Epstein, R. A. (2011). Distances between real-world locations are represented in the human hippocampus. *The Journal of Neuroscience*, 31(4):1238–1245.

Morris, A. M., Churchwell, J. C., Kesner, R. P., and Gilbert, P. E. (2012). Selective lesions of the dentate gyrus produce disruptions in place learning for adjacent spatial locations. *Neurobiology of Learning and Memory*, 97:326–331.

Motley, S. E. and Kirwan, C. B. (2012). A parametric investigation of pattern separation processes in the medial temporal lobe. *Journal of Neuroscience*, 33(28):13076–13084.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, 89(6):609–626.

Murdock, B. B. (1995). Similarity in a distributed memory model. *Journal of Mathematical Psychology*, 39(3):251–264.

Naselaris, T., Kay, K. N., Nishimoto, S., and Gallant, J. L. (2011). Encoding and decoding in fMRI. *NeuroImage*, 56(2):400–410.

Nasr, S., Echavarria, C. E., and Tootell, R. B. H. (2014). Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. *The Journal of Neuroscience*, 34(20):6721–6735.

Nasr, S. and Tootell, R. B. H. (2012). A cardinal orientation bias in scene-selective visual cortex. *The Journal of Neuroscience*, 32(43):14921–14926.

Navawongse, R. and Eichenbaum, H. (2013). Distinct pathways for rule-based retrieval and spatial mapping of memory representations in hippocampal neurons. *The Journal of Neuroscience*, 33(3):1002–1013.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5):1170–1187.

Naveh-Benjamin, M., Guez, J., Kilb, A., and Reedy, S. (2004). The associative memory deficit of older adults: further support using face-name associations. *Psychology and Aging*, 19(3):541–546.

Neunuebel, J. P. and Knierim, J. J. (2014). CA3 receives coherent representations from degraded input: direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron*, 81:416–427.

Norman, K. A. (2010). How hippocampus and cortex contribute to recognition memory: revisiting the Complementary Learning Systems model. *Hippocampus*, 20:1217–1227.

Norman, K. A. and O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological Review*, 110(4):611–646.

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9):242–430.

Okado, Y. and Stark, C. E. L. (2005). Neural activity during encoding predicts false memories created by misinformation. *Learning and Memory*, 12:3–11.

O'Keefe, J. and Dostrovsky, J. (1971). The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1):171–175.

Old, S. R. and Naveh-Benjamin, M. (2008a). Differential effects of age on item and associative measures of memory: a meta-analysis. *Psychology and Aging*, 23(1):104–118.

Old, S. R. and Naveh-Benjamin, M. (2008b). Memory for people and their actions: further evidence for an age-related associative deficit. *Psychology and Aging*, 23(2):467–472.

Oliva, A. and Torralba, A. (2001). Modeling the shape of the scene: a holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42(3):145–175.

Op de Beeck, H. P., Vermaercke, B., Woolley, D. G., and Wenderoth, N. (2013). Combinatorial brain decoding of people's whereabouts during visuospatial navigation. *Frontiers in Neuroscience*, 7:78.

O'Reilly, R. C. and McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage and recall: avoiding a tradeoff. *Hippocampus*, 4(6):661–682.

O'Reilly, R. C. and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. The MIT Press.

O'Reilly, R. C. and Norman, K. A. (2002). Hippocampal and neocortical contributions to memory: advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, 6(12):505–510.

O'Reilly, R. C. and Rudy, J. W. (2000). Computational principles of learning in the neocortex and hippocampus. *Hippocampus*, 10:389–397.

O'Reilly, R. C. and Rudy, J. W. (2001). Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychological Review*, 108(2):311–345.

Otten, L. J., Henson, R. N. A., and Rugg, M. D. (2001). Depth of processing effects on neural correlates of memory encoding. *Brain*, 124:399–412.

Packard, M. G. and McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory*, 65:65–72.

Pais-Vieira, M., Lebedev, M., Kunicki, C., Wang, J., and Nicolelis, M. A. L. (2013). A brain-to-brain interface for real-time sharing of sensorimotor information. *Scientific Reports*, 3:1319.

Park, S. and Chun, M. M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *NeuroImage*, 47:1747–1756.

Parron, C. and Save, E. (2004). Comparison of the effects of entorhinal and retrosplenial cortical lesions on habituation, reaction to spatial and non-spatial changes during object exploration in the rat. *Neurobiology of Learning and Memory*, 82:1–11.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(437–442).

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., and Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, 331:585–589.

Polyn, S. M., Natu, V. S., Cohen, J. D., and Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science*, 310(5756):1963–1966.

Potvin, O., Doré, F. Y., and Goulet, S. (2009). Lesions of the dorsal subiculum and the dorsal hippocampus impaired pattern separation in a task using distinct and overlapping visual stimuli. *Neurobiology of Learning and Memory*, 91:287–297.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, 59(3):2142–2154.

Rajimehr, R., Devaney, K. J., Bilenko, N. Y., Young, J. C., and Tootell, R. B. H. (2011). The "parahippocampal place area" responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biology*, 9(4).

Rajji, T., Chapman, D., Eichenbaum, H., and Greene, R. (2006). The role of CA3 hippocampal NMDA receptors in paired associate learning. *The Journal of Neuroscience*, 26(3):908–915.

Ranganath, C. (2010). A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus*, 20:1263–1290.

Ranganath, C. and Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience*, 13:713–726.

Reagh, Z. M. and Yassa, M. A. (2014). Repitition strengthens target recognition but impairs similar lure discrimination: evidence for trace competition. *Learning & Memory*, 21(7):342–346.

Reed, J. M. and Squire, L. R. (1997). Impaired recognition memory in patients with lesions limited to the hippocampal formation. *Behavioral Neuroscience*, 13:667–675.

Rey, A. (1941). Lexamen psychologique dans les cas dencephalopathie traumatique. *Archives of Psychology*, 28:286–340.

Rissman, J., Gazzaley, A., and D'Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, 23:752–763.

Ritchey, M., Libby, L. A., and Ranganath, C. (2015). Cortico-hippocampal systems involved in memory and cognition: the PMAT framework. *Progress in Brain Research*, 219:45–64.

Ritchey, M., Yonelinas, A. P., and Ranganath, C. (2014). Functional connectivity relationships predict similarities in task activation and pattern information during associative memory encoding. *Journal of Cognitive Neuroscience*, 26(5):1085–1099.

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology*, 53(2):94–101.

Rugg, M. D., Fletcher, P. C., Frith, C. D., Frackowiak, R. S. J., and Dolan, R. J. (1997). Brain regions supporting intentional and incidental memory: a PET study. *Neuroreport*, 8:1283–1287.

Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*. The MIT Press.

Rumelhart, D. E. and Todd, P. M. (1993). Learning and connectionist representations. In Meyers, D. E. and Kornblum, S., editors, *Attention and Performance XIV*, pages 3–30. MIT Press/Bradford Books.

Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., and Cox, R. W. (2009). A new method for improving functional-to-structural alignment using local Pearson correlation. *NeuroImage*, 44:839–848.

Saksida, L. M. and Bussey, T. J. (2010). The representational-hierarchical view of amnesia: translation from animal to human. *Neuropsychologia*, 48:2370–2384.

Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., and Ranganath, C. (2015). Memory and space: towards an understanding of the cognitive map. *The Journal of Neuroscience*, 35(41):13904–13911.

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

Shepard, R. N. (1987). Perceptual-cognitive universals as reflections of the world. *Science*, 237:1317–1323.

Shepard, R. N. and Chipman, S. (1970). Second-order isomorphism of internal representations: shapes of states. *Cognitive Psychology*, 1:1–17.

Sherry, D. F. and Schacter, D. L. (1987). The evolution of multiple memory system. *Psychological Review*, 94(4):439–454.

Silson, E. H., Chan, A. W. Y., Reynolds, R. C., Kravitz, D. J., and Baker, C. I. (2015). A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *The Journal of Neuroscience*, 35(34):11921–11935.

Skaggs, W. E. and McNaughton, B. L. (1998). Spatial firing properties of hippocampal CA1 populations in an environment containing two visually identical regions. *The Journal of Neuroscience*, 18(20):8455–8466.

Smith, D. G. and Duncan, M. J. J. (2004). Testing theories of recognition memory by predicting performance across paradigms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(3):615–625.

Smith, D. M. and Bulkin, D. A. (2014). The form and function of hippocampal context representations. *Neuroscience and Biobehavioral Reviews*, 40:52–61.

Smith, D. M. and Mizumori, S. J. Y. (2006a). Hippocampal place cells, context, and episodic memory. *Hippocampus*, 16(9):716–729.

Smith, D. M. and Mizumori, S. J. Y. (2006b). Learning-related development of context-specific neuronal responses to places and events: the hippocampal role in context processing. *The Journal of Neuroscience*, 26(12):3154–3163.

Spencer, W. D. and Raz, N. (1995). Differential effects of aging on memory for content and context: a meta-analysis. *Psychology and Aging*, 10(4):527–539.

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99(3):195–231.

Squire, L. R., Wixted, J. T., and Clark, R. E. (2007). Recognition memory and the medial temporal lobe: a new perspective. *Nature Reviews Neuroscience*, 8:872–883.

Squire, L. R. and Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences USA*, 93:13515–13522.

Stanislaw, H. and Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1):137–149.

Stark, C. E. L. and Okado, Y. (2003). Making memories without trying: medial temporal lobe activity associated with incidental memory formation during recognition. *The Journal of Neuroscience*, 23(17):6748–6753.

Stark, C. E. L. and Squire, L. R. (2001). When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of the Sciences USA*, 98(22):12760–12766.

Stark, C. E. L. and Squire, L. R. (2003). Hippocampal damage equally impairs memory for single items and memory for conjunctions. *Hippocampus*, 13:281–292.

Stark, S. M., Stevenson, R., Wu, C., Rutledge, S., and Stark, C. E. L. (2015). Stability of age-related deficits in the Mnemonic Similarity Task across task variations. *Behavioral Neuroscience*, 129(3):257–268.

Stark, S. M., Yassa, M. A., Lacy, J. W., and Stark, C. E. L. (2013). A task to assess behavioral pattern separation (BPS) in humans: data from healthy aging and mild cognitive impairment. *Neuropsychologia*, 51(12):2442–2449.

Stelzer, J., Chen, Y., and Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. *NeuroImage*, 65:69–82.

Suzuki, W. A. (2009). Perception and the medial temporal lobe: evaluating the current evidence. *Neuron*, 61:657–666.

Suzuki, W. A. and Amaral, D. G. (1994). Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *Journal of Comparative Neurology*, 350:497–533.

Swets, J. A. and Pickett, R. M. (1982). *Evaluation of diagnostic systems*. Academic Press.

Talpos, J. C., McTighe, S. M., Dias, R., Saksida, L. M., and Bussey, T. J. (2010). Trial-unique, delayed nonmatching-to-location (TUNL): A novel, highly hippocampus-dependent automated touchscreen test of location memory and pattern separation. *Neurobiology of Learning and Memory*, 94:341–352.

Tan, R. H., Wong, S., Hodges, J. R., Halliday, G. M., and Hornberger, M. (2013). Retrosplenial cortex (BA 29) volumes in behavioral variant frontotemporal dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 35:177–182.

Tanaka, K. Z., Pevzner, A., Hamidi, A. B., Nakazawa, Y., Graham, J., and Wiltgen, B. J. (2014). Cortical representations are reinstated by the hippocampus during memory retrieval. *Neuron*, 84(2):347–354.

Teyler, T. J. and DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral Neuroscience*, 100(2):147–154.

Teyler, T. J. and Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus*, 17:1158–1169.

Theriault, C., Thome, N., and Cord, M. (2011). HMAX-S: deep scale. *ICIP*, pages 1261–1264.

Tonor, C. K., Pirogovsky, E., Kirwan, C. B., and Gilbert, P. E. (2009). Visual object pattern separation deficits in nondemented older adults. *Learning and Memory*, 16:338–342.

Tort, A. B. L., Komorowski, R., Kopell, N., and Eichenbaum, H. (2013). A mechanism for the formation of hippocampal neuronal firing patterns that represent what happens where. *Learning & Memory*, 18(11):718–727.

Treves, A. and Rolls, E. T. (1992). Computational constraints suggest the need for two distinct input systems to the hippocampal CA3 network. *Hippocampus*, 2:189–199.

Treves, A. and Rolls, E. T. (1994). Computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4:374–391.

Tulving, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning and Verbal Behavior*, 20(5):479–496.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26(1):1–12.

Turner, B. O., Mumford, J. A., Poldrack, R. A., and Ashby, F. G. (2012). Spatiotemporal activity estimation for multivoxel pattern analysis with rapid event-related designs. *NeuroImage*, 62(3):1429–1438.

Ungerleider, L. G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science*, 270(5237):769–775.

Valenstein, E., Bowers, D., Verfaellie, M., Heilman, K. M., Day, A., and Watson, R. T. (1987). Retrosplenial amnesia. *Brain*, 110:1631–1646.

Van Dijk, K. R. A., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity. *NeuroImage*, 59(1):431–438.

Vann, S. D. and Aggleton, J. P. (2002). Extensive cytotoxic lesions of the rat retrosplenial cortex reveal consistent deficits on tasks that tax allocentric spatial memory. *Behavioral Neuroscience*, 116(1):85–94.

Vann, S. D., Aggleton, J. P., and Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, 10:792–802.

Vass, L. K. and Epstein, R. A. (2013). Abstract representations of location and facing direction in the human brain. *The Journal of Neuroscience*, 33(14):6133–6142.

Vass, L. K. and Epstein, R. A. (2016). Common neural representations for visually guided reorientation and spatial imagery. *Cerebral Cortex*, doi: 10.1093/cercor/bhv343.

Vazdarjanova, A. and Guzowski, J. F. (2004). Differences in hippocampal neuronal population responses to modification of an environmental context: evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles. *The Journal of Neuroscience*, 24(29):6489–6496.

Villain, N., Desgranges, B., Viader, F., de la Sayette, V., Mézenge, F., Landeau, B., Baron, J. C., Eustache, F., and Chételat, G. (2008). Relationships between hippocampal atrophy, white matter disruption, and gray matter hypometabolism in Alzheimer's disease. *The Journal of Neuroscience*, 28(24):6174–6181.

Wais, P. E., Squire, L. R., and Wixted, J. T. (2010). In search of recollection and familiarity in the hippocampus. *Journal of Cognitive Neuroscience*, 22:109–123.

Walther, D. B., Beck, D. M., and Fei-Fei, L. (2012). To err is human: correlating fMRI decoding and behavioral errors to probe the neural representation of natural scene categories. In Kriegeskorte, N. and Kreimen, G., editors, *Visual Population Codes*, pages 391–415. The MIT Press.

Walther, D. B., Caddigan, E., Fei-Fei, L., and Beck, D. M. (2009). Natural scene categories revealed in distributed patterns of activity in the human brain. *The Journal of Neuroscience*, 29:10573–10581.

Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience*, 5(4):279–290.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G., Gosselin, F., and Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavioral Research Methods*, 42(3):671–684.

Wilson, I. A., Gallagher, M., Eichenbaum, H., and Tanila, H. (2006). Neurocognitive aging: prior memories hinder new hippocampal encoding. *Trends in Neurosciences*, 29(12):662–670.

Wing, E. A., Ritchey, M., and 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.

Wixted, J. T., Mickes, L., and Squire, L. R. (2010). Measuring recollection and familiarity in the medial temporal lobe. *Hippocampus*, 20:1195–1205.

Wixted, J. T. and Squire, L. R. (2011). The medial temporal lobe and the attributes of memory. *Trends in Cognitive Sciences*, 15(5):210–217.

Wood, E. R., Dudchenko, P. A., Robitsek, R. J., and Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occuring in the same location. *Neuron*, 27(3):623–633.

Wright, P., Randall, B., Clarke, A., and Tyler, L. K. (2015). The perirhinal cortex and conceptual processing: effects of feature-based statistics following damage to the anterior temporal lobes. *Neuropsychologia*, 76:192–207.

Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M., and Stark, C. E. (2011a). Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*, 21(9):968–979.

Yassa, M. A., Mattfeld, A. T., Stark, S. M., and Stark, C. E. L. (2011b). Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proceedings of the National Academy of Sciences USA*, 108(21):8873–8878.

Yassa, M. A. and Stark, C. E. L. (2009). A quantitative evaluation of cross-participant alignment techniques for MRI studies of the medial temporal lobe. *NeuroImage*, 44(2):319–327.

Yassa, M. A. and Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, 34(10):515–525.

Yonelinas, A. P. (1994). Receiver operating characteristics in recognition memory: evidence for a dual process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6):1341–1354.

Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: the contribution of recollection and familiarity. *Memory and Cognition*, 25(6):747–763.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *Journal of Memory and Language*, 36:441–517.

Yonelinas, A. P., Aly, M., Wang, W. C., and Koen, J. D. (2010). Recollection and familiarity: examining controversial assumptions and new directions. *Hippocampus*, 20:1178–1194.

Zhang, S. and Manahan-Vaughan, D. (2015). Spatial olfactory learning contributes to place field formation in the hippocampus. *Cerebral Cortex*, 25(2):423–432.