Representational Structures of Overlapping Memories: From Behavior to Brain

by

Anisha S. Babu

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Psychology

Dissertation Committee:

Brice A. Kuhl, Chair

Dagmar (Dasa) Zeithamova, Core Member

J. Benjamin Hutchinson, Core Member

James M. Murray, Institutional Representative

University of Oregon

Summer 2026

© 2026 Anisha S. Babu
This work is openly licensed via Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0 International license.



DISSERTATION ABSTRACT

Anisha S. Babu

Doctor of Philosophy in Psychology

Title: Representational Structures of Overlapping Memories: From Behavior to Brain

Ongoing experience comprises highly overlapping events, and this overlap shapes memory encoding. While the influence of overlap on memory has been extensively studied, prior studies are largely limited to controlled stimuli and simple measures of memory. As such, this dissertation explores how overlap influences memory in more naturalistic contexts. Specifically, across three empirical chapters, I use naturalistic scene stimuli and verbal recall descriptions to consider how overlap shapes: (1) how overlapping events are remembered relative to one another, and (2) how neural encoding of overlapping events relates to idiosyncratic versus shared features of subsequent memory.

Across two experiments in **Chapter II**, I explore how verbal recall is influenced by overlapping stimuli. In the first experiment, I found that overlap leads to a separation in verbal recall for multiple overlapping stimuli – analogous with prior accounts of repulsion of overlapping memories. In the second experiment, I considered the effect of overlap using pairs of overlapping stimuli and found that the specific overlapping competitor is associated with a consistent but nondirectional effect on verbal recall. In **Chapter III**, I develop a new method to analyze data from the second experiment in **Chapter II**. Specifically, I used image segmentation to consider whether verbal recall prioritizes specific 'patches' of an image based on the specific competing image. I found that different patches of an image are indeed prioritized for different competing images. Moreover, patch importance corresponded with 'objective' dissimilarity between the pairs of overlapping images at that patch. These two chapters use innovative methods to demonstrate dynamic influences of overlap on verbal recall of naturalistic scene images at the behavioral level. In **Chapter IV**, I consider how neural representations during encoding of overlapping events relate to idiosyncratic versus shared features in subsequent verbal recall. I found that visual and parietal cortex dissociably reflect idiosyncratic and shared features in verbal recall, respectively. Moreover, I found that hippocampus negatively reflects idiosyncratic verbal recall – consistent with hippocampal differentiation of overlapping memories. Collectively, these chapters provide insight into the subtle and multidimensional ways in which overlap influences behavioral and neural expressions of episodic memory in naturalistic contexts.

This dissertation includes unpublished co-authored material.

ACKNOWLEDGMENTS

I would like to thank all faculty and staff in the psychology department, the Institute of Neuroscience, and the Lewis Center for Neuroimaging at the University of Oregon. Their time, support, and instruction have been invaluable during my graduate education. Additionally, I owe thanks to our graduate employee union (GTFF), whose labor allowed me to feel represented and supported during my graduate studies. I am also indebted to the NIH for funding via an NRSA F31 fellowship. Finally, I would like to thank to the hundreds of individuals who contributed their time to participate in my research, and the research assistants who made this data collection possible.

I am very grateful to my advisors for their guidance and support throughout my education. During my undergraduate, Dr. Julie Golomb provided an extremely positive and fulfilling undergraduate research experience, through which I developed a passion for cognitive neuroscience. During my PhD, Dr. Brice Kuhl has been a great advisor and sponsor whose thoughtful and insightful guidance encourages me to think more deeply and creatively about my research. I am very grateful for his humor and kindness as he supported me over the last six years. I would like to thank my cosponsor Dr. James Murray, whose unique perspectives and feedback helped me approach my research using interdisciplinary methods. Finally, I thank the remainder of my dissertation committee, Dr. Dasa Zeithamova and Dr. Ben Hutchinson, for their instrumental time, support, and feedback.

I would like to thank my parents for their love and support over the years. Their steadfast presence, care, and homecooked meals have given me the strength to accomplish my goals from the other side of the country. I am grateful for all the opportunities (and extra homework) they provided me growing up, without which I could not have discovered my current path. My dad, Suresh, inspired me to work hard and dream big, and always pushed me to be the best version of myself. My mom, Anu, has always believed in me and reminded me of the bright possibilities for the future.

Lastly, I thank my friends and family (near and far) whose support, companionship, and humor have carried me through this period of my life. I can always count on my sister, Varsha, to entertain me with silly stories and pictures of my niece-cat Kajol. I am grateful to my peers in the 2020 (COVID) cohort, without whom the transition to graduate school would have been very isolating. I am also grateful for the amazing friends I have found within the Kuhl lab, with whom

I have made great memories eating good food, exploring Oregon, and distracting each other at work.

For Ajesh.

Thank you for being my biggest supporter – my friend, my family, my love. I treasure the memories we made together and will carry you forever in my heart.

TABLE OF CONTENTS

Chapter Page

I. INTRODUCTION 15

Effects of overlap on memory 16

Behavioral 16

Neuroimaging 17

Overlap and idiosyncrasies in memory 19

Behavioral 19

Neuroimaging 19

Approaches for defining overlap 20

Overview of the present work 21

References Cited 22

II. STIMULUS OVERLAP SHAPES VERBAL RECALL OF NATURALISTIC SCENES 31

Introduction 31

Experiment 1 33

Methods 33

Participants 33

Stimuli 34

Experimental procedure 34

Vividness ratings 36

Verbal descriptions 36

Verbal recall 38

Overlap-driven differences in verbal recall 38
Results 39
Successful encoding across learning task 39
Verbal descriptions 40
Verbal recall 40
Overlap-driven differences in verbal recall 40
Experiment 2 43
Methods 43
Participants 43
Stimuli 44
Experimental procedure 45
Verbal recall 46
Specific overlap-driven differences in verbal recall 47
Results 48
Successful encoding across learning task 48
Image descriptions as a measure of memory content 48
Specific overlap-driven differences in verbal recall 49
General Discussion 50
References Cited 54

Chapter Page

III. VERBAL RECALL OF OVERLAPPING SCENES PRIORITIZES DISSIMILAR IMAGE PATCHES 58

Introduction 58

Methods 59

Experimental design 59

Image segmentation 60

Recall importance map analysis 61

Patch importance based on VGG-16 similarity 63

Results 64

Verbal recall prioritizes competitor-specific image patches 64

Patch importance driven by VGG-16 similarity 72

General Discussion 74

References Cited 75

IV. NEURAL ENCODING PREDICTS SHARED VERSUS IDIOSYNCRATIC VERBAL RECALL 78

Introduction 78

Methods 80

Participants 80

Stimuli 80

Experimental procedure 81

Analysis of behavioral data 83

Vividness ratings 83

Verbal recall 84
Representational similarity analysis 84
Category sensitivity in recall 85
Overlap with VGG-16 85
Shared representational structure in recall 85
fMRI acquisition 86
fMRI preprocessing 86
Regions of interest 88
Analysis of neural data 89
Representational similarity analysis 89
Category sensitivity in neural activity 90
Shared representational structure in neural activity 90
Modeling neural representations of recall 90
Results 91
Behavioral measures of recall 91
fMRI measures of representational structure during encoding 93
Relating fMRI structures of encoding structure to text-based measures of recall structure 95
General Discussion 100
References Cited 101
V. GENERAL DISCUSSION 112
Quantifying complex visual memories 113
Overlap shapes memory for naturalistic scene images 114

Conclusion 116

References Cited 117

LIST OF FIGURES

Figure Page

2.1. Stimuli & task design for Experiment 1 36

2.2. Organization of recall in Experiment 1. 40

2.3. Experiment 1 results 41

2.4. Organization of descriptions in annotation task 44

2.5. Stimuli & task design for Experiment 2 46

2.6. Experiment 2 results 49

3.1. Recall importance map analysis 61

3.2. Pool recall importance maps. 64

3.3. Skating recall importance maps..... 65

3.4. Library recall importance maps 66

3.5. Pond recall importance maps..... 67

3.6. Stadium recall importance maps 68

3.7. Downtown recall importance maps 69

3.8. Competitor-specific recall importance maps 70

3.9. VGG-16 similarity predicts recall importance 72

4.1. Stimuli & task design 82

4.2. Regions of interest. 88

4.3. Verbal recall results. 93

4.4. Neural mixed effects model results 98

4.5. Supplementary permutation results 99

LIST OF TABLES

Table Page

4.1. Statistics from Model 1 (three VGG-16 layer predictors) 95

4.2. Statistics from Model 2 (participant-specific recall and VGG-16 layer 3 predictors) 96

4.3. Statistics from Model 3 (participant-specific recall and normative recall predictors) 97

CHAPTER I
INTRODUCTION

Ongoing experience comprises both highly variable and highly overlapping events, which shape episodic memory. For example, an individual may experience very unique visits to Disney World and Tokyo Disney but very similar visits to Disney World and Disneyland. The extent of overlap between these visits can influence how these visits are subsequently remembered. Moreover, an individual's prior experiences can shape how new events are subjectively experienced. For example, a Disney aficionado's subjective memory of visiting Tokyo Disney would likely be much different from that of a Disney novice. As illustrated, overlap can shape episodic memory in a myriad of ways including: (1) how overlapping events are remembered relative to one another, and (2) how overlapping events are encoded in the context of an individual's idiosyncratic background.

It has been well documented that overlap can influence memory success – in particular by determining *whether* an event is remembered and *how* it is remembered. In some contexts, overlap can improve memory (Antony et al., 2022; Staresina et al., 2009), while in other contexts, overlap can lead to interference and forgetting (M. C. Anderson & Spellman, 1995; Mensink & Raaijmakers, 1988; Osgood, 1949). Meanwhile, overlap can also affect how an event is remembered relative to other events, with prior studies revealing both attraction and repulsion effects between overlapping memories (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024), depending on the nature of a task and task goals (Chanales et al., 2021; Lively et al., 2021; Melega & Sheldon, 2023; Robertson, 2018). Moreover, neuroimaging studies demonstrate that overlap influences neural representations across cortex and hippocampus. On one hand, cortical regions represent overlapping visual and semantic features (Clarke & Tyler, 2014; Davis et al., 2020; Devereux et al., 2018; Seghier, 2013), which can be shaped by top-down attention to different features (Chapman et al., 2023; Kok, Brouwer, Gerven, et al., 2013; Maunsell & Treue, 2006). The hippocampus, on the other hand, actively integrates and adaptively differentiates overlapping memories (Chanales et al., 2017; Favila et al., 2016; Hulbert & Norman, 2015; Kim et al., 2017; Kuhl et al., 2010; LaRocque et al., 2013; Libby et al., 2019; Mirjalili et al., 2025; V. J. Ritvo et al., 2024; Schlichting et al., 2014, 2015; Varga et al., 2025; Wanjia et al., 2021; Wing et al., 2020). However, it remains unclear whether differences in how overlapping events are remembered relate to how they are represented in the brain.

Most approaches to studying the effects of overlap on memory assume that overlap acts in a consistent manner across individuals. Indeed, if memories are encoded purely based on perceptual input, individuals receiving identical perceptual input should form identical memories. However, different individuals possess unique knowledge and experiences, which may interact with and shape how new experiences are encoded into memory. As such, individuals experiencing the same event often form different, idiosyncratic, memories of that event (Bartlett, 1932; Chen et al., 2017; Heusser et al., 2021). However, most studies consider neural representations of idiosyncratic memory as dissimilarity *between* individuals, therefore rendering idiosyncratic differences indistinguishable from measurement noise (Finn et al., 2020; Hasson et al., 2004; Nastase et al., 2019). In order to consider idiosyncrasies *within* individuals, new methods are required that allow for idiosyncratic and shared representations to be considered simultaneously. Given these limitations, the relationship between overlap and idiosyncratic memory content has remained largely unexplored.

The aim of this dissertation is to consider the various ways in which overlap influences episodic memory. In a series of behavioral and fMRI experiments, I use verbal recall as a measure of memory content in order to explore how overlap affects memory content for overlapping experiences, and how overlapping experiences are encoded within an individual's idiosyncratic context. The main goals include: (1) investigating how overlap affects verbal recall for naturalistic scene stimuli, (2) developing a method to identify image feature preferences in verbal recall of overlapping scenes, and (3) determining how idiosyncratic verbal recall for overlapping stimuli is represented in the brain. The remainder of this chapter will review relevant literature in the field before describing each chapter in more detail.

Effects of overlap on memory

Behavioral

Overlap can influence memory success in different ways. In some situations, overlap between events can lead to better memory. Specifically, new experiences may (1) scaffold to and strengthen older related memories (i.e., retroactive facilitation; Antony et al., 2022) or (2) improve memory by increasing semantic elaboration (Staresina et al., 2009). On the other hand, overlap can lead to memory interference and forgetting (Anderson & Spellman, 1995; Mensink & Raaijmakers, 1988; Osgood, 1949). In these situations, certain memories are prioritized at the

expense of other memories. For example, older memories may disrupt formation of memory for new overlapping experiences (proactive interference; Wickens, 1973; Wickens et al., 1963) or new overlapping experiences can disrupt older memories (retroactive interference; (Baddeley & Dale, 1966; Bower et al., 1994; McGeoch & McDonald, 1931). Moreover, retrieval of certain memories can lead to forgetting of other overlapping memories (retrieval-induced forgetting; Anderson et al., 1994, 2000). This foundational body of work largely focuses on memory outcomes in terms of memory success (i.e., whether a stimulus was remembered or not). However, overlap can also affect *how* a stimulus is remembered.

Recent studies have begun to explore how overlap affects memory for specific stimulus features. These studies demonstrate that overlap can lead to distortions in memory of a stimulus' features (e.g., color, orientation, etc.; Bae & Luck, 2017; Chanales et al., 2021; Chapman et al., 2023; Chunharas et al., 2022; Drascher & Kuhl, 2022; Golomb, 2015; Morrow & Clewett, 2025; Rademaker et al., 2015; Tandoc et al., 2024; Teng & Kravitz, 2019). While it is tempting to consider these distortions as a failure of memory, in some cases, distortions may be adaptive in nature. Indeed, several episodic memory studies have observed distortion of overlapping memories toward each other (i.e., attraction; Chanales et al., 2021; Morrow & Clewett, 2025; Tandoc et al., 2024), especially when it is beneficial to make inferences across these overlapping memories (Chanales et al., 2021). Conversely, when discrimination is essential to task goals, it may be adaptive to distort memories away from each other. Visual working memory studies have shown that memory for color (Chapman et al., 2023; Chunharas et al., 2022; Golomb, 2015) or orientation (Bae & Luck, 2017) of an object can be shifted away from that of an overlapping distractor. Meanwhile, episodic memory studies observe a similar change where memory for color (Chanales et al., 2021) or facial features (Drascher & Kuhl, 2022) are adaptively separated, resulting in reduced memory interference. Nevertheless, these studies largely focus on memory for a single controlled stimulus feature, making it unclear whether (and how) memory might be distorted in more complex and naturalistic contexts.

Neuroimaging

Humans constantly receive an enormous amount of perceptual input through ongoing experience, which must somehow be represented within the brain. Along the visual processing stream, visual and temporal cortex represent low-level and high-level visual features, respectively

(Clarke & Tyler, 2014; Davis et al., 2020; Devereux et al., 2018). Meanwhile, parietal cortex is thought to represent more conceptual and semantic information (Davis et al., 2020; Seghier, 2013). Taken together, different regions in cortex appear to play different roles in representing features (i.e., visual and semantic) of overlapping experiences. However, a large body of work indicates that even these fundamental feature representations play a role in preventing memory interference between overlapping experiences. For example, Naspi et al. (2021) and Morales-Torres et al. (2024) show that fine-grained representations in visual cortex are especially important for memory success. Moreover, several studies have demonstrated that stronger category-specific (Kuhl et al., 2012) and item-specific (Feng et al., 2019; Kok, Brouwer, Gerven, et al., 2013; Lu et al., 2015; Wagner et al., 2016; Ward et al., 2013; Xue et al., 2010, 2013; Zheng et al., 2018) representations across repeated encoding of the same event are associated with better memory.

While much of cortex represents basic visual and semantic features of a given event (Clarke & Tyler, 2014; Davis et al., 2020; Devereux et al., 2018; Seghier, 2013), the hippocampus is a key region involved in adaptively representing overlapping experiences based on prior experience and current goals (O'Reilly & McClelland, 1994). For example, when encoding a new experience, the hippocampus may reactivate prior overlapping experiences. This reactivation can impair memory when prior information is prioritized over new information (Sherman & Turk-Browne, 2020). However, in other cases, reactivation can lead to improved memory (Kuhl et al., 2010; Schlichting et al., 2014), possibly by facilitating integration of overlapping memories across cortex (Chanales et al., 2019; Koen & Rugg, 2016; Molitor et al., 2021; Peng et al., n.d.; Shohamy & Wagner, 2008; Zeithamova et al., 2012). Conversely, hippocampus also plays a role in adaptively separating representations of overlapping memories (Chanales et al., 2017; Favila et al., 2016; Hulbert & Norman, 2015; Kim et al., 2017; Mirjalili et al., 2025; Wanjia et al., 2021; Wing et al., 2020) in order to overcome memory interference and improve memory outcomes (Favila et al., 2016; Hulbert & Norman, 2015; Wanjia et al., 2021; Wing et al., 2020). These complementary integration and differentiation processes are thought to scale with the degree of stimulus overlap (V. J. H. Ritvo et al., 2019; Wammes et al., 2022) and be prompted by different contexts and task goals (Brunec et al., 2020; Fernandez et al., 2023; Libby et al., 2019; V. J. Ritvo et al., 2024; Schlichting et al., 2015; Varga et al., 2025). However, it remains unclear whether these adaptive changes in hippocampal representations coincide with adaptive changes in memory content.

Overlap and idiosyncrasies in memory

Behavioral

Overlap is known to affect memory via interaction with prior experiences and knowledge. Overlapping prior experiences (Antony et al., 2022) and knowledge (Bein et al., 2015) can improve memory, as experiences that are congruent with an individual's understanding of the world are more likely to be remembered. However, prior experience and knowledge can be highly variable across individuals, which may lead to idiosyncratic differences in memory when individuals experience the same event (Bartlett, 1932; Chen et al., 2017; Heusser et al., 2021). Therefore, an important consideration is how overlap relates to idiosyncratic differences in memory. On one hand, overlap may not be associated with idiosyncrasies across individuals, instead affecting memory in a generalizable manner. Indeed, many studies have shown convergence of narrative memory across individuals (Bartlett, 1932; Bower et al., 1979; Chen et al., 2017), and even prioritization of the same items in memory (i.e., memorability; Bainbridge et al., 2017). On the other hand, an individual's prior experiences may bias memory in an idiosyncratic manner (Robson & Tangen, 2023). Indeed, when perceptual information is largely overlapping for overlapping experiences, it may be more subjective interpretations that help distinguish overlapping events. Moreover, idiosyncratic attention to different details or spatial locations may lead to encoding of different, idiosyncratic, combinations of features in memory. However, limited work has explored true idiosyncratic memory due to methodological limitations; differences in memory success may simply reflect fluctuations in attention (deBettencourt et al., 2018; Wakeland-Hart et al., 2022). Moreover, it is difficult to isolate idiosyncratic memory from shared memory, as individuals will likely possess a high degree of shared information in memory even with idiosyncratic influences.

Neuroimaging

Idiosyncratic information in memory must be represented in the brain during encoding, as neural representations during encoding determine what can later be retrieved. However, neural representations of idiosyncratic memory possess their own measurement difficulties. Specifically, neuroimaging studies of idiosyncratic differences have largely considered idiosyncratic neural activity patterns to be the inverse of shared representations (i.e., the less similar activity patterns are across individuals, the more idiosyncratic they are; Finn et al., 2020; Hasson et al., 2004; Nastase et al., 2019). However, as discussed in Finn et al. (2020), across-subject measurements of

idiosyncratic activity patterns are impossible to discriminate from measurement noise, as differences across individuals may simply reflect fluctuations in attention (i.e., noise) rather than true idiosyncratic differences (Ki et al., 2016; Song et al., 2021). Given these limitations, it becomes even more difficult to isolate how overlap influences idiosyncratic memory representations. Perhaps overlap leads to increased top-down influence during encoding (Kok, Brouwer, van Gerven, et al., 2013; Maunsell & Treue, 2006), therefore increasing the amount of idiosyncratic neural representations. Alternatively, overlap may amplify idiosyncratic variability in the specific features that are attended to and, consequently, encoded. Either way, it is possible that adaptive changes in neural representations of overlapping experiences rely partially on idiosyncratic features in memory. Yet, due to methodological constraints, the dynamics between idiosyncratic neural representations and overlapping experiences have remained largely unexplored.

Approaches for defining overlap

A key consideration when studying the effects of overlap on memory content and neural representations is how to quantify overlap between stimuli. As shown in prior studies, the degree of initial overlap determines the nature of effect on memory. Specifically, moderate overlap leads to hippocampal differentiation (Wammes et al., 2022) and repulsion in memory (Chanales et al., 2021) while high overlap leads to integration of memories. These findings provide support for the nonmonotonic plasticity hypothesis (NMPH), which states that the initial coactivation (i.e., overlap) between two memories determines whether those memories will later be weakened (therefore eliciting adaptive separation processes) or strengthened (Lewis-Peacock & Norman, 2014; Newman & Norman, 2010; V. J. H. Ritvo et al., 2019). However, the degree of stimulus overlap can be quantified in many different ways. Early studies have defined overlap between word stimuli as subjective semantic relatedness (Roediger & McDermott, 1995). More recently, word stimulus overlap has been quantified using large language models such as word2vec (Gatti et al., 2023; Walsh & Rissman, 2023). Meanwhile, studies using image stimuli have largely used highly controlled stimuli that vary along one or two easily quantifiable dimensions (e.g., color or facial features; Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024; Wammes et al., 2022). However, newer computer vision models such as VGG-16 (Simonyan & Zisserman, 2015) afford an alternate method to define image overlap, allowing for

the use of more naturalistic real-world images while still controlling for stimulus overlap (Groen et al., 2018; Mirjalili et al., 2025; Morales-Torres et al., 2024).

While quantifying stimulus overlap during experimental design is important, an equally important consideration is quantifying stimulus overlap behaviorally. Simple measures of establishing stimulus overlap include pairwise similarity judgments, perceptual discrimination tasks, and similarity arrangement tasks (Daggett & Hout, 2025; Hout et al., 2013; Jiang et al., 2022; Migo et al., 2013; Son et al., 2022). Recently, naturalistic methods have provided more complex behavioral measures of overlap. For example, eye-tracking can be used to assess gaze patterns for different stimuli (N. C. Anderson et al., 2015; klein Selle et al., 2021). Meanwhile drawings (Bainbridge et al., 2019) and verbal descriptions (Fenerci et al., 2025; Schooler & Engstler-Schooler, 1990; Wang et al., 2023) can be used as measures of visual and semantic overlap between images. Verbal descriptions are particularly well-suited when considering overlap between images, as large language models can also be used to quantify descriptions in multidimensional space. In summary, advances in naturalistic behavioral methods allow for images to be quantified with a high level of detail, which will be essential in considering how overlap and idiosyncratic differences subtly influence memory content.

Overview of the present work

The present work adds to the literature described above by using verbal recall as a measure of memory for naturalistic scene images, and assessing how overlap and idiosyncratic differences influence behavioral and neural expressions of memory content. Across three empirical studies, I will show that overlap influences memory content in verbal recall (**Chapters II and III**) and develop a new method to isolate image feature preferences in verbal recall of scene images (**Chapter III**). Additionally, I will demonstrate that different brain regions variably reflect shared and idiosyncratic verbal recall for overlapping stimuli (**Chapter IV**). A key method in each of these studies is the use of verbal recall as a rich and complex measure of behavioral memory content. Verbal recall possesses a unique profile of features which may be quantified and studied rigorously using large language models. By using verbal recall, this body of work adds to our understanding of how overlap influences episodic memory in more naturalistic contexts.

In **Chapter II**, I explore whether overlap shapes verbal recall for naturalistic scene images. In a first experiment, I found that overlap from multiple competing images leads to separation in

verbal recall. This finding is consistent with either (1) targeted separation, or (2) a general response to overlap. In a second experiment, I alternated which overlapping competitor image participants studied to examine whether the separation in recall is targeted based on the specific competitor image. I found that the specific overlapping competitor influences verbal recall in a consistent but nondirectional manner. This finding provides insight into the subtle influences of overlap on verbal recall which cannot be fully explained by traditional accounts of attraction or repulsion.

In **Chapter III**, I introduce a new method to isolate image feature preferences in verbal recall using data from the second experiment of Chapter II. Specifically, I used image segmentation to identify the specific patches of a scene that are most important in verbal recall, and generated recall importance maps for a given image based on its specific overlapping competitor. I found that, after optimizing patch resolution, importance maps were consistent across participants with the same overlapping competitor. Moreover, importance of a given patch corresponded with the degree of objective dissimilarity between overlapping stimuli at that location – suggesting verbal recall prioritizes features in diagnostic parts of overlapping scene images.

In **Chapter IV** (coauthored material by Anisha S. Babu, Zhifang Ye, Julian Gamez, and Brice A. Kuhl), I consider whether fMRI activity patterns reflect participant-specific and normative verbal recall of overlapping stimuli as a proxy of idiosyncratic and shared verbal recall, respectively. I find a dissociation across cortical regions wherein visual cortex better reflects participant-specific (i.e., idiosyncratic) recall while parietal cortex reflects both participant-specific and normative (i.e., shared) recall. I also provide evidence for a negative relationship between participant-specific recall and hippocampal representations, consistent with hippocampal differentiation (Chanales et al., 2017; Favila et al., 2016; Hulbert & Norman, 2015; Kim et al., 2017; Mirjalili et al., 2025; Wanjia et al., 2021; Wing et al., 2020). **Chapter V** will summarize these findings and situate them within the broader field.

References Cited

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review*, *7*(3), 522–530. https://doi.org/10.3758/bf03214366

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1063–1087. https://doi.org/10.1037/0278-7393.20.5.1063

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*(1), 68–100. https://doi.org/10.1037/0033-295X.102.1.68

Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2015). A comparison of scanpath comparison methods. *Behavior Research Methods*, *47*(4), 1377–1392. https://doi.org/10.3758/s13428-014-0550-3

Antony, J. W., Romero, A., Vierra, A. H., Luenser, R. S., Hawkins, R. D., & Bennion, K. A. (2022). Semantic relatedness retroactively boosts memory and promotes memory interdependence across episodes. *eLife*, *11*, e72519. https://doi.org/10.7554/eLife.72519

Baddeley, A. D., & Dale, H. C. A. (1966). The effect of semantic similarity on retroactive interference in long- and short-term memory. *Journal of Verbal Learning & Verbal Behavior*, *5*(5), 417–420. https://doi.org/10.1016/S0022-5371(66)80054-3

Bae, G.-Y., & Luck, S. J. (2017). Interactions between visual working memory representations. *Attention, Perception & Psychophysics*, *79*(8), 2376–2395. https://doi.org/10.3758/s13414-017-1404-8

Bainbridge, W. A., Dilks, D. D., & Oliva, A. (2017). Memorability: A stimulus-driven perceptual neural signature distinctive from memory. *NeuroImage*, *149*, 141–152. https://doi.org/10.1016/j.neuroimage.2017.01.063

Bainbridge, W. A., Hall, E. H., & Baker, C. I. (2019). Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. *Nature Communications*, *10*(1), 5. https://doi.org/10.1038/s41467-018-07830-6

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology* (pp. xix, 317). Cambridge University Press.

Bein, O., Livneh, N., Reggev, N., Gilead, M., Goshen-Gottstein, Y., & Maril, A. (2015). Delineating the Effect of Semantic Congruency on Episodic Memory: The Role of Integration and Relatedness. *PLOS ONE*, *10*(2), e0115624. https://doi.org/10.1371/journal.pone.0115624

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*(2), 177–220. https://doi.org/10.1016/0010-0285(79)90009-4

Bower, G. H., Thompson-Schill, S., & Tulving, E. (1994). Reducing retroactive interference: An interference analysis. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(1), 51–66. https://doi.org/10.1037//0278-7393.20.1.51

Brunec, I. K., Robin, J., Olsen, R. K., Moscovitch, M., & Barense, M. D. (2020). Integration and differentiation of hippocampal memory traces. *Neuroscience & Biobehavioral Reviews*, *118*, 196–208. https://doi.org/10.1016/j.neubiorev.2020.07.024

Chanales, A. J. H., Dudukovic, N. M., Richter, F. R., & Kuhl, B. A. (2019). Interference between overlapping memories is predicted by neural states during learning. *Nature Communications*, *10*(1), 5363. https://doi.org/10.1038/s41467-019-13377-x

Chanales, A. J. H., Oza, A., Favila, S. E., & Kuhl, B. A. (2017). Overlap among Spatial Memories Triggers Repulsion of Hippocampal Representations. *Current Biology*, *27*(15), 2307-2317.e5. https://doi.org/10.1016/j.cub.2017.06.057

Chanales, A. J. H., Tremblay-McGaw, A. G., Drascher, M. L., & Kuhl, B. A. (2021). Adaptive Repulsion of Long-Term Memory Representations Is Triggered by Event Similarity. *Psychological Science*, *32*(5), 705–720. https://doi.org/10.1177/0956797620972490

Chapman, A. F., Chunharas, C., & Störmer, V. S. (2023). Feature-based attention warps the perception of visual features. *Scientific Reports*, *13*(1), 6487. https://doi.org/10.1038/s41598-023-33488-2

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Chunharas, C., Rademaker, R. L., Brady, T. F., & Serences, J. T. (2022). An adaptive perspective on visual working memory distortions. *Journal of Experimental Psychology. General*, *151*(10), 2300–2323. https://doi.org/10.1037/xge0001191

Clarke, A., & Tyler, L. K. (2014). Object-specific semantic coding in human perirhinal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(14), 4766–4775. https://doi.org/10.1523/JNEUROSCI.2828-13.2014

Daggett, E. W., & Hout, M. C. (2025). A tutorial review on methods for collecting similarity judgments from human observers. *Attention, Perception & Psychophysics*, *87*(3), 737–751. https://doi.org/10.3758/s13414-025-03044-3

Davis, S. W., Geib, B. R., Wing, E. A., Wang, W.-C., Hovhannisyan, M., Monge, Z. A., & Cabeza, R. (2020). Visual and Semantic Representations Predict Subsequent Memory in Perceptual and Conceptual Memory Tests. *Cerebral Cortex (New York, NY)*, *31*(2), 974–992. https://doi.org/10.1093/cercor/bhaa269

deBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin & Review*, *25*(2), 605–611. https://doi.org/10.3758/s13423-017-1309-5

Devereux, B. J., Clarke, A., & Tyler, L. K. (2018). Integrated deep visual and semantic attractor neural networks predict fMRI pattern-information along the ventral object processing pathway. *Scientific Reports*, *8*(1), 10636. https://doi.org/10.1038/s41598-018-28865-1

Drascher, M. L., & Kuhl, B. A. (2022). Long-term memory interference is resolved via repulsion and precision along diagnostic memory dimensions. *Psychonomic Bulletin & Review*, *29*(5), 1898–1912. https://doi.org/10.3758/s13423-022-02082-4

Favila, S. E., Chanales, A. J. H., & Kuhl, B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nature Communications*, *7*(1), 11066. https://doi.org/10.1038/ncomms11066

Fenerci, C., Cheng, Z., Addis, D. R., Bellana, B., & Sheldon, S. (2025). Studying memory narratives with natural language processing. *Trends in Cognitive Sciences*, *29*(6), 516–525. https://doi.org/10.1016/j.tics.2025.02.003

Feng, K., Zhao, X., Liu, J., Cai, Y., Ye, Z., Chen, C., & Xue, G. (2019). Spaced Learning Enhances Episodic Memory by Increasing Neural Pattern Similarity Across Repetitions. *The Journal of Neuroscience*, *39*(27), 5351–5360. https://doi.org/10.1523/JNEUROSCI.2741-18.2019

Fernandez, C., Jiang, J., Wang, S.-F., Choi, H. L., & Wagner, A. D. (2023). Representational integration and differentiation in the human hippocampus following goal-directed navigation. *eLife*, *12*, e80281. https://doi.org/10.7554/eLife.80281

Finn, E. S., Glerean, E., Khojandi, A. Y., Nielson, D., Molfese, P. J., Handwerker, D. A., & Bandettini, P. A. (2020). Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *NeuroImage*, *215*, 116828. https://doi.org/10.1016/j.neuroimage.2020.116828

Gatti, D., Marelli, M., Mazzoni, G., Vecchi, T., & Rinaldi, L. (2023). Hands-on false memories: A combined study with distributional semantics and mouse-tracking. *Psychological Research*, *87*(4), 1129–1142. https://doi.org/10.1007/s00426-022-01710-x

Golomb, J. D. (2015). Divided spatial attention and feature-mixing errors. *Attention, Perception, & Psychophysics*, *77*(8), 2562–2569. https://doi.org/10.3758/s13414-015-0951-0

Groen, I. I., Greene, M. R., Baldassano, C., Fei-Fei, L., Beck, D. M., & Baker, C. I. (2018). Distinct contributions of functional and deep neural network features to representational similarity of scenes in human brain and behavior. *eLife*, *7*, e32962. https://doi.org/10.7554/eLife.32962

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Heusser, A. C., Fitzpatrick, P. C., & Manning, J. R. (2021). Geometric models reveal behavioural and neural signatures of transforming experiences into memories. *Nature Human Behaviour*, *5*(7), 905–919. https://doi.org/10.1038/s41562-021-01051-6

Hout, M. C., Goldinger, S. D., & Ferguson, R. W. (2013). The versatility of SpAM: A fast, efficient, spatial method of data collection for multidimensional scaling. *Journal of Experimental Psychology: General*, *142*(1), 256–281. https://doi.org/10.1037/a0028860

Hulbert, J. C., & Norman, K. A. (2015). Neural Differentiation Tracks Improved Recall of Competing Memories Following Interleaved Study and Retrieval Practice. *Cerebral Cortex*, *25*(10), 3994–4008. https://doi.org/10.1093/cercor/bhu284

Jiang, Z., Sanders, D. M. W., & Cowell, R. A. (2022). Visual and semantic similarity norms for a photographic image stimulus set containing recognizable objects, animals and scenes. *Behavior Research Methods*, *54*(5), 2364–2380. https://doi.org/10.3758/s13428-021-01732-0

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention Strongly Modulates Reliability of Neural Responses to Naturalistic Narrative Stimuli. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(10), 3092–3101. https://doi.org/10.1523/JNEUROSCI.2942-15.2016

Kim, G., Norman, K. A., & Turk-Browne, N. B. (2017). Neural Differentiation of Incorrectly Predicted Memories. *The Journal of Neuroscience*, *37*(8), 2022–2031. https://doi.org/10.1523/JNEUROSCI.3272-16.2017

klein Selle, N., Gamer, M., & Pertzov, Y. (2021). Gaze-pattern similarity at encoding may interfere with future memory. *Scientific Reports*, *11*(1), 7697. https://doi.org/10.1038/s41598-021-87258-z

Koen, J. D., & Rugg, M. D. (2016). Memory Reactivation Predicts Resistance to Retroactive Interference: Evidence from Multivariate Classification and Pattern Similarity Analyses. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(15), 4389–4399. https://doi.org/10.1523/JNEUROSCI.4099-15.2016

Kok, P., Brouwer, G. J., Gerven, M. A. J. van, & Lange, F. P. de. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *Journal of Neuroscience*, *33*(41), 16275–16284. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Kok, P., Brouwer, G. J., van Gerven, M. A. J., & de Lange, F. P. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *The Journal of Neuroscience*, *33*(41), 16275–16284. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

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

Kuhl, B. A., Shah, A. T., DuBrow, S., & Wagner, A. D. (2010). Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. *Nature Neuroscience*, *13*(4), 501–506. https://doi.org/10.1038/nn.2498

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

Lewis-Peacock, J. A., & Norman, K. A. (2014). Competition between items in working memory leads to forgetting. *Nature Communications*, *5*(1), 5768. https://doi.org/10.1038/ncomms6768

Libby, L. A., Reagh, Z. M., Bouffard, N. R., Ragland, J. D., & Ranganath, C. (2019). The hippocampus generalizes across memories that share item and context information. *Journal of Cognitive Neuroscience*, *31*(1), 24–35. https://doi.org/10.1162/jocn_a_01345

Lively, Z., Robinson, M. M., & Benjamin, A. S. (2021). Memory Fidelity Reveals Qualitative Changes in Interactions Between Items in Visual Working Memory. *Psychological Science*, *32*(9), 1426–1441. https://doi.org/10.1177/0956797621997367

Lu, Y., Wang, C., Chen, C., & Xue, G. (2015). Spatiotemporal Neural Pattern Similarity Supports Episodic Memory. *Current Biology*, *25*(6), 780–785. https://doi.org/10.1016/j.cub.2015.01.055

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. https://doi.org/10.1016/j.tins.2006.04.001

McGeoch, J. A., & McDonald, W. T. (1931). Meaningful Relation and Retroactive Inhibition. *The American Journal of Psychology*, *43*(4), 579–588. https://doi.org/10.2307/1415159

Melega, G., & Sheldon, S. (2023). Conceptual relatedness promotes memory generalization at the cost of detailed recollection. *Scientific Reports*, *13*(1), 15575. https://doi.org/10.1038/s41598-023-40803-4

Mensink, G.-J., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review*, *95*(4), 434–455. https://doi.org/10.1037/0033-295X.95.4.434

Migo, E. M., Montaldi, D., & Mayes, A. R. (2013). A visual object stimulus database with standardized similarity information. *Behavior Research Methods*, *45*(2), 344–354. https://doi.org/10.3758/s13428-012-0255-4

Mirjalili, S., Wanjia, G., Grätz, D., Wang, E., Mayr, U., & Kuhl, B. A. (2025). Hippocampal transformations occur along dimensions of memory interference. *bioRxiv*, 2025.10.13.682242. https://doi.org/10.1101/2025.10.13.682242

Molitor, R. J., Sherrill, K. R., Morton, N. W., Miller, A. A., & Preston, A. R. (2021). Memory Reactivation during Learning Simultaneously Promotes Dentate Gyrus/CA2,3 Pattern Differentiation and CA1 Memory Integration. *Journal of Neuroscience*, *41*(4), 726–738. https://doi.org/10.1523/JNEUROSCI.0394-20.2020

Morales-Torres, R., Wing, E. A., Deng, L., Davis, S. W., & Cabeza, R. (2024). Visual Recognition Memory of Scenes Is Driven by Categorical, Not Sensory, Visual Representations. *Journal of Neuroscience*, *44*(21). https://doi.org/10.1523/JNEUROSCI.1479-23.2024

Morrow, E., & Clewett, D. (2025). Distortion of overlapping memories relates to arousal and anxiety. *Cognitive, Affective & Behavioral Neuroscience*, *25*(1), 154–172. https://doi.org/10.3758/s13415-024-01229-1

Naspi, L., Hoffman, P., Devereux, B., & Morcom, A. M. (2021). Perceptual and Semantic Representations at Encoding Contribute to True and False Recognition of Objects. *Journal of Neuroscience*, *41*(40), 8375–8389. https://doi.org/10.1523/JNEUROSCI.0677-21.2021

Nastase, S. A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience*, *14*(6), 667–685. https://doi.org/10.1093/scan/nsz037

Newman, E. L., & Norman, K. A. (2010). Moderate Excitation Leads to Weakening of Perceptual Representations. *Cerebral Cortex (New York, NY)*, *20*(11), 2760–2770. https://doi.org/10.1093/cercor/bhq021

O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, *4*(6), 661–682. https://doi.org/10.1002/hipo.450040605

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review*, *56*(3), 132–143. https://doi.org/10.1037/h0057488

Peng, K., Wammes, J. D., Nguyen, A., Iordan, C. R., Norman, K. A., & Turk-Browne, N. B. (n.d.). Inducing representational change in the hippocampus through real-time neurofeedback. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1915), 20230091. https://doi.org/10.1098/rstb.2023.0091

Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(6), 1650–1665. https://doi.org/10.1037/xhp0000110

Ritvo, V. J. H., Turk-Browne, N. B., & Norman, K. A. (2019). Nonmonotonic Plasticity: How Memory Retrieval Drives Learning. *Trends in Cognitive Sciences*, *23*(9), 726–742. https://doi.org/10.1016/j.tics.2019.06.007

Ritvo, V. J., Nguyen, A., Turk-Browne, N. B., & Norman, K. A. (2024). A neural network model of differentiation and integration of competing memories. *eLife*, *12*, RP88608. https://doi.org/10.7554/eLife.88608

Robertson, E. M. (2018). Memory instability as a gateway to generalization. *PLOS Biology*, *16*(3), e2004633. https://doi.org/10.1371/journal.pbio.2004633

Robson, S. G., & Tangen, J. M. (2023). The invisible 800-pound gorilla: Expertise can increase inattentional blindness. *Cognitive Research: Principles and Implications*, *8*, 33. https://doi.org/10.1186/s41235-023-00486-x

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*(1), 8151. https://doi.org/10.1038/ncomms9151

Schlichting, M. L., Zeithamova, D., & Preston, A. R. (2014). CA1 subfield contributions to memory integration and inference. *Hippocampus*, *24*(10), 1248–1260. https://doi.org/10.1002/hipo.22310

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, *22*(1), 36–71. https://doi.org/10.1016/0010-0285(90)90003-M

Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *19*(1), 43–61. https://doi.org/10.1177/1073858412440596

Sherman, B. E., & Turk-Browne, N. B. (2020). Statistical prediction of the future impairs episodic encoding of the present. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(37), 22760–22770. https://doi.org/10.1073/pnas.2013291117

Shohamy, D., & Wagner, A. D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, *60*(2), 378–389. https://doi.org/10.1016/j.neuron.2008.09.023

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv:1409.1556). arXiv. https://doi.org/10.48550/arXiv.1409.1556

Son, G., Walther, D. B., & Mack, M. L. (2022). Scene wheels: Measuring perception and memory of real-world scenes with a continuous stimulus space. *Behavior Research Methods*, *54*(1), 444–456. https://doi.org/10.3758/s13428-021-01630-5

Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences*, *118*(33), e2021905118. https://doi.org/10.1073/pnas.2021905118

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event Congruency Enhances Episodic Memory Encoding through Semantic Elaboration and Relational Binding. *Cerebral Cortex (New York, NY)*, *19*(5), 1198–1207. https://doi.org/10.1093/cercor/bhn165

Tandoc, M. C., Dong, C. V., & Schapiro, A. C. (2024). Object Feature Memory Is Distorted by Category Structure. *Open Mind: Discoveries in Cognitive Science*, *8*, 1348–1368. https://doi.org/10.1162/opmi_a_00170

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour*, *3*(8), 827–836. https://doi.org/10.1038/s41562-019-0640-4

Varga, N. L., Roome, H. E., Molitor, R. J., Martinez, L., Hipskind, E. M., Mack, M. L., Preston, A. R., & Schlichting, M. L. (2025). Differentiation of Related Events in Hippocampus Supports Memory Reinstatement in Development. *Journal of Cognitive Neuroscience*, *37*(4), 853–894. https://doi.org/10.1162/jocn_a_02299

Wagner, I. C., van Buuren, M., Bovy, L., & Fernández, G. (2016). Parallel Engagement of Regions Associated with Encoding and Later Retrieval Forms Durable Memories. *The Journal of Neuroscience*, *36*(30), 7985–7995. https://doi.org/10.1523/JNEUROSCI.0830-16.2016

Wakeland-Hart, C. D., Cao, S. A., deBettencourt, M. T., Bainbridge, W. A., & Rosenberg, M. D. (2022). Predicting visual memory across images and within individuals. *Cognition*, *227*, 105201. https://doi.org/10.1016/j.cognition.2022.105201

Walsh, C., & Rissman, J. (2023). *Episodic learning-induced representational change is both influenced by and re-sculpts semantic memory* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/8kq59

Wammes, J., Norman, K. A., & Turk-Browne, N. (2022). Increasing stimulus similarity drives nonmonotonic representational change in hippocampus. *eLife*, *11*, e68344. https://doi.org/10.7554/eLife.68344

Wang, Y., Lee, H., & Kuhl, B. A. (2023). *Mapping multidimensional content representations to neural and behavioral expressions of episodic memory* [Preprint]. Neuroscience. https://doi.org/10.1101/2023.01.09.523096

Wanjia, G., Favila, S. E., Kim, G., Molitor, R. J., & Kuhl, B. A. (2021). Abrupt hippocampal remapping signals resolution of memory interference. *Nature Communications*, *12*(1), 4816. https://doi.org/10.1038/s41467-021-25126-0

Ward, E. J., Chun, M. M., & Kuhl, B. A. (2013). Repetition Suppression and Multi-Voxel Pattern Similarity Differentially Track Implicit and Explicit Visual Memory. *Journal of Neuroscience*, *33*(37), 14749–14757. https://doi.org/10.1523/JNEUROSCI.4889-12.2013

Wickens, D. D. (1973). Some characteristics of word encoding. *Memory & Cognition*, *1*(4), 485–490. https://doi.org/10.3758/BF03208913

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *2*(5), 440–445. https://doi.org/10.1016/S0022-5371(63)80045-6

Wing, E. A., Geib, B. R., Wang, W.-C., Monge, Z., Davis, S. W., & Cabeza, R. (2020). Cortical Overlap and Cortical-Hippocampal Interactions Predict Subsequent True and False Memory. *Journal of Neuroscience*, *40*(9), 1920–1930. https://doi.org/10.1523/JNEUROSCI.1766-19.2020

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

Xue, G., Dong, Q., Chen, C., Lu, Z.-L., Mumford, J. A., & Poldrack, R. A. (2013). Complementary role of frontoparietal activity and cortical pattern similarity in successful episodic memory encoding. *Cerebral Cortex*, *23*(7), 1562–1571. https://doi.org/10.1093/cercor/bhs143

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, *75*(1), 168–179. https://doi.org/10.1016/j.neuron.2012.05.010

Zheng, L., Gao, Z., Xiao, X., Ye, Z., Chen, C., & Xue, G. (2018). Reduced Fidelity of Neural Representation Underlies Episodic Memory Decline in Normal Aging. *Cerebral Cortex*, *28*(7), 2283–2296. https://doi.org/10.1093/cercor/bhx130

CHAPTER II

STIMULUS OVERLAP SHAPES VERBAL RECALL OF NATURALISTIC SCENES

Introduction

As we navigate the world, we often experience highly overlapping events, which affects our ability to successfully encode those events into memory (Anderson & Spellman, 1995; Antony et al., 2022; Mensink & Raaijmakers, 1988; Osgood, 1949; Staresina et al., 2009). In some cases, overlap can improve memory (Antony et al., 2022; Staresina et al., 2009), while in other cases it can cause memory interference, reducing the likelihood that an event is remembered (Anderson & Spellman, 1995; Mensink & Raaijmakers, 1988; Osgood, 1949). More recent work indicates that even when an event is successfully remembered, overlap can influence *how* that event is remembered relative to other overlapping events. Specifically, behavioral studies have shown that overlapping experiences can sometimes be remembered as more similar to one another (i.e., attraction; Chanales et al., 2021; Morrow & Clewett, 2025; Tandoc et al., 2024). On the other hand, some studies reveal an opposite effect: features of overlapping stimuli become separated in memory (i.e., repulsion; Chanales et al., 2021; Drascher & Kuhl, 2022; Zhao et al., 2021). fMRI studies have found analogous results, where hippocampal activity patterns for overlapping stimuli become separated (i.e., hippocampal differentiation; Dimsdale-Zucker et al., 2018; Favila et al., 2016; Hulbert & Norman, 2015; Schlichting et al., 2015; Wanjia et al., 2021). Interestingly, this separation of overlapping experiences is shown to be adaptive, wherein separation (both behavioral and fMRI) is associated with improved memory for those experiences (Chanales et al., 2021; Drascher & Kuhl, 2022; Favila et al., 2016; Hulbert & Norman, 2015; Wanjia et al., 2021). However, these studies have largely used highly constrained stimuli and behavioral measures of memory. As such, a critical consideration is whether and how overlap influences memory for overlapping experiences in more naturalistic contexts.

It is possible that subtle memory effects will not be observable using more naturalistic measures. Indeed, prior work observing behavioral repulsion has only used pairs of overlapping stimuli that vary along one or two specific feature dimensions (e.g., color or facial features; Chanales et al., 2021; Drascher & Kuhl, 2022; Zhao et al., 2021). Real-world experiences are comprised of extremely rich and complex information, which cannot be captured within a single dimension. Moreover, by using simple artificially generated stimuli, prior studies were also constrained to simple behavioral tests (e.g., multiple-choice recognition or explicit recall of one or

two features). As such, to improve generalizability to real-world memories, it is necessary to consider more naturalistic stimuli and behavioral measures. With more naturalistic stimuli, one possibility is that increased richness of information circumvents the need to adaptively discriminate overlapping memories. On the other hand, if overlap does still influence memory for more naturalistic stimuli, it remains uncertain whether established behavioral measures of memory would be sensitive enough to detect these changes; subtle differences may be masked by high dimensionality of stimuli and behavioral measures.

To date, many studies have used naturalistic stimuli to study the effects of overlap on memory, including naturalistic scene images (Chanales et al., 2017, 2021; Morales-Torres et al., 2024; Wanjia et al., 2021), movies (Chadwick et al., 2011; Hebscher et al., 2023; Reagh & Ranganath, 2023), and narratives (Cohn-Sheehy et al., 2021). However, these studies typically consider simple behavioral measures of memory such as recognition (Chanales et al., 2017; Morales-Torres et al., 2024; Wanjia et al., 2021) or explicit recall of specific details (Chanales et al., 2017, 2021), which lack sufficient detail to capture the richness of naturalistic stimuli. A promising alternative, which is both unconstrained and includes extensive detail, is verbal recall descriptions. Though verbal recall descriptions have been used to capture information in memories for scene images (Wang et al., 2023), movies (Chen et al., 2017), and narratives (Cohn-Sheehy et al., 2021), they have not been used to compare differences in memory for overlapping events. In order to quantify and compare differences in verbal descriptions, large language models are uniquely well-suited (Wang et al., 2023). Large language models are trained on an extensive corpus of text and, when given a new text input, output a vector of hundreds of numbers representing semantic information in the text (i.e., a text embedding). Using these models, verbal recall descriptions can be used to consider subtle differences in memory for complex overlapping naturalistic stimuli.

In the current chapter, I report findings from two behavioral experiments assessing how overlap influences memory. In both experiments, participants first studied naturalistic scene images, then described them from memory. We later quantified verbal recall descriptions using a large language model. In Experiment 1, we utilized a between-subjects design to assess differences in verbal recall between participants that experienced overlap and participants that experienced no overlap. We hypothesized that lower similarity in verbal recall among participants in the overlap group would be indicative of a separation in verbal recall due to overlap (analogous to behavioral

repulsion). Results from Experiment 1 indicate overlap indeed leads to lower similarity in verbal recall. However, since participants studied multiple overlapping stimuli, it remains unclear whether this separation in recall reflects a general response to overlap versus targeted repulsion based on the specific studied images. Therefore, in Experiment 2, we included only two overlapping scene images to assess whether the specific studied images drive differences in verbal recall.

Experiment 1

Methods

In Experiment 1, we alternately recruited subjects into two groups: overlap and nonoverlap. Participants in the overlap group studied a set of six scene images drawn from a single semantic category. Participants in the nonoverlap group studied a set of six scene images drawn from six *different* semantic categories. In order to perform secondary analyses, we collected data in an annotation task where an independent set of participants provided verbal descriptions of images they were actively viewing as a baseline measure of features in each image.

Participants

For Experiment 1 and the annotation task, target sample sizes were identified in advance, and additional participants were recruited to meet the target after applying exclusion criteria. To reach these samples, participants were recruited from the University of Oregon subject pool (Experiment 1) and online through Prolific (annotation task). All participants self-reported to be right-handed, native English speakers, and have normal or corrected-to-normal vision. Informed consent was obtained from all participants according to a protocol approved by the University of Oregon Institutional Review Board, and all participants received credit (Experiment 1) or were paid $10/hour (annotation task) for their participation. Because we did not have a way to estimate effect sizes in advance, we chose a target sample of 120 participants for Experiment 1 (i.e., 60 participants in the overlap and nonoverlap groups; 83 female; mean age 19.11 ± 1.57 SD, range 18-30). For the annotation task, we chose a target sample of 225 participants (133 female; mean age 27.10 ± 4.51 SD, range 18-35). Exclusion criteria are described in the Experimental Procedure section.

Stimuli

The annotation task used a set of 90 scene images comprised of 15 images drawn equally from each of six categories: indoor pool, ice skating rink, library, pond, stadium, and downtown street. Participants were recruited into one of 15 image subset groups. Each subset group was assigned one image from each of the six categories, with no image overlap between groups. Participants were evenly distributed across the 15 groups (i.e., 15 participants per group). All participants were recruited online through Prolific to describe their subset of six scene images. Participants were shown each scene image at the top of the screen and were asked to type a description below. Participants were provided with a good example description in the instructions. After typing their response, they pressed 'continue' to advance to the next trial. A time limit of 60s was provided, and participants were required to use at least 15 words in their descriptions to proceed (or the trial timed out). To ensure engagement with the task, participants were informed that at least half of their responses must have at least 15 words to receive payment for participation. Participants missing descriptions and/or possessing descriptions with <10 words were excluded from analysis (n=46). Additional participants were recruited to make up for these participants to maintain balanced participants for each image subset group.

For Experiment 1, 36 scene images and six face images were used. The 36 scene images included in Experiment 1 were drawn from the 90 images used in the annotation task, with six images drawn from each of the six categories (Fig. 2.1a). Each participant was assigned six images to study and later describe from memory. Each of the six scene images assigned to a participant was randomly paired with a face image, creating a set of six face-scene paired associates for each participant. The faces served as cues during cued recall tasks (described below).

Experimental procedure

In Experiment 1, participants were alternately recruited into either the overlap or nonoverlap groups. In the overlap group, participants were assigned all six scene images from a single category (e.g., all six indoor pool images). Participants were evenly distributed across the six categories (i.e., 10 participants per category). In the nonoverlap group, participants were assigned a single image from each of the six categories. As with the overlap group, participants were evenly distributed such that each image was described by 10 different participants.

Participants completed six runs of a learning task (Fig. 2.1b). Each run began with a self-initiated start screen and included two phases: encoding and retrieval. The start of each phase was preceded by a screen introducing the phase as "Study" or "Test" (i.e., encoding or retrieval; 5500ms). Across all phases, stimuli were displayed on a white background. After completing the learning task, participants completed two additional memory tasks. The experiment was implemented in PsychoPy (Peirce et al., 2019) and jsPsych (Leeuw et al., 2023) hosted on Pavlovia (https://pavlovia.org), and lasted approximately 1.5 hours.

During the encoding phases, participants learned all face-scene associations, one association at a time. Phases were separated into two blocks, with each association presented once within each block. Each encoding trial began with the presentation of a face image (1000ms), followed by a black fixation cross (500ms), the associated scene image (3000ms), then another black fixation cross (1000ms). The order in which the face-scene associations were encoded was randomized within each run and block, and for each participant.

During the retrieval phases, participants were cued to recall the scene associated with each of the faces, one association at a time. Phases were separated into two blocks, with each association cued once within each block. Each retrieval trial began with a face image (1000ms), followed by a black fixation cross (500ms), then a white box (1500ms) as a cue to retrieve the associated scene image. When the white box turned red (1500ms), participants were asked to report vividness of their memory via key press on a three-point scale ("no memory", "weak memory", "vivid memory"). A black fixation cross (1000ms) was then presented until the start of the next trial. The order in which the face-scene associations were retrieved was randomized within each run and block, and for each participant.

After completing the learning task, participants completed a cued recall test. For each face-scene associate, participants were shown a cue (face) at the top of the screen and were asked to type a description of the associated scene image (Fig. 2.1c). Specifically, participants were asked to complete a description beginning with: "An image of …" which was already displayed on the screen. Further, they were asked to avoid describing the face image, using unnecessary words (e.g., restating "an image of"), or referencing other scene images. Example good and bad descriptions were included in the instructions. After participants typed their response, they pressed 'continue' to advance to the next trial. No time limit was given, though participants were required to use at least 10 words in their descriptions to proceed. We chose to reduce the minimum number of words

from 15 words in the annotation task due to the relative higher difficulty of recall. If participants did not remember the associated scene image, they could check a box indicating that they did not remember and were forced to wait 30s before proceeding to the next trial.

Following cued verbal recall, participants completed a multiple-choice recognition test for each of the face-scene associates. Here, participants were presented with each face image at the top of the screen and six scene images below. The six scene images were the same ones that a given participant had studied. Participants were asked to click on the correct scene image using the mouse, then click on their confidence (sure/guess). The recognition test was separated into two blocks, with each face presented once within each block. The order in which the scene images were tested was randomized within each block, and for each participant. Participants missing any verbal recall response and/or incorrectly responding to more than one multiple-choice question were excluded from analysis (n=12). Additional participants were recruited to make up for these participants to maintain balanced participants for each group.

Vividness ratings

In each of the six retrieval phases, participants rated the vividness of their memory for each scene image on a 3-point scale ("no memory", "weak memory", "vivid memory"). Vividness ratings were recoded to a numerical scale (1-3). To compare ratings across runs, we applied a mixed effects model using the lme4 package in R to predict mean vividness rating as a function of learning run with participant as a random effect to account for individual differences: vividness_rating ~ learning_run + (1|participant). We also ran Welch's two-sample t-tests to assess differences in initial and final vividness ratings based on overlap and nonoverlap groups.

Verbal descriptions

In the annotation task, participants described each scene image they were presented with. To determine if there were any effects of image category on the length of verbal descriptions, we compared the mean number of words for each category using a repeated measures ANOVA with factor category.

Next, we defined features in verbal descriptions by transforming descriptions into text embeddings using the large language model MPNet. We removed punctuation, extra whitespace, and changed all characters to lowercase using the Gensim preprocessing library (Rehurek, 2011).

We then transformed each description into a text embedding using MPNet (i.e., 'paraphrase-mpnet-base-v2' from the Sentence Transformers library; Reimers & Gurevych, 2019; Song et al., 2020). This yielded 768-dimensional vectors representing semantic information in verbal descriptions for each image and each participant.

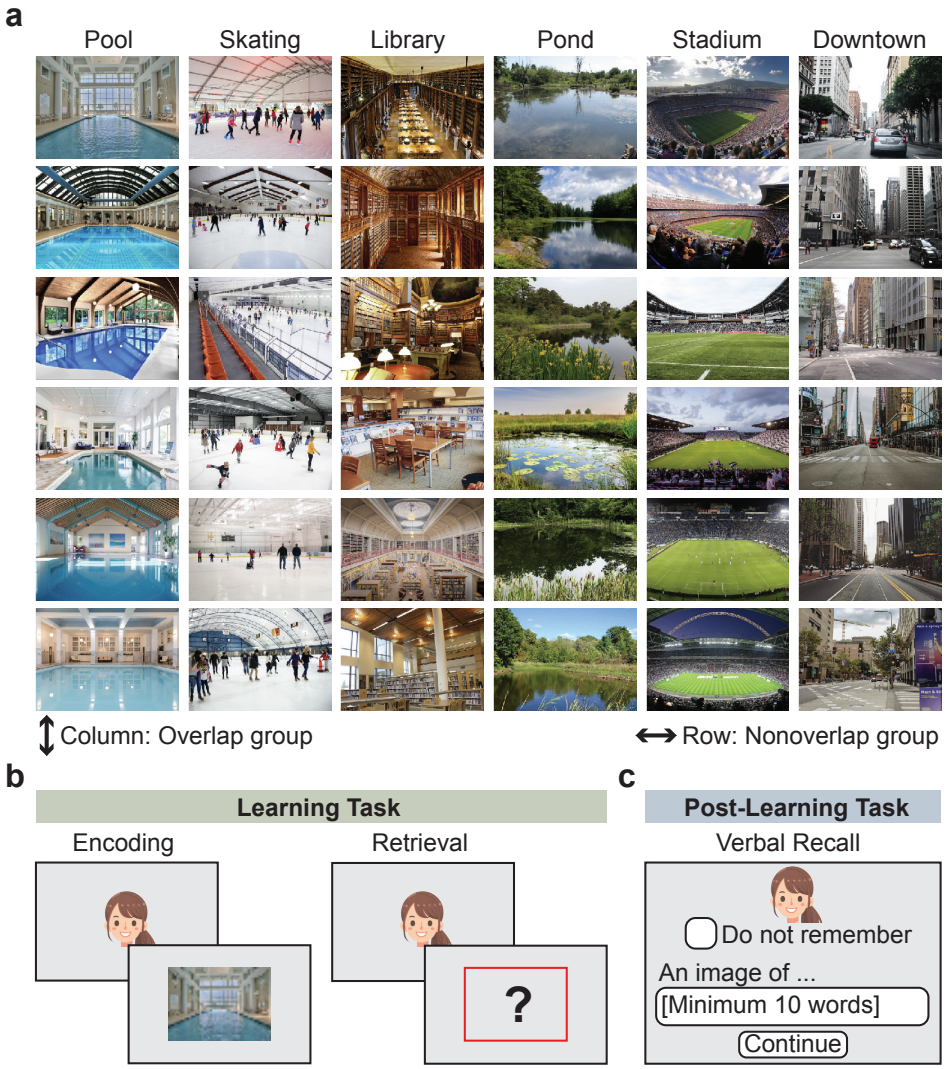

Figure 2.1. Stimuli & task design for Experiment 1. a) Stimuli. Scene images used in Experiment 1. Participants in the overlap group were assigned a single category of images (e.g., column). Participants in the nonoverlap group were assigned one image from each of six categories (e.g., row). **b)** Learning task. Learning consisted of encoding and retrieval trials. **c)** Recall task. After learning, participants were cued with each face image at the top of the screen and asked to type a description of the associated scene image.

Text embeddings from participants in the annotation task were used to represent a baseline measure of features in each image. As such, we calculated an average text embedding for each of the 90 images. These text embeddings were used for analyses in Experiment 1 and stimulus selection in Experiment 2 (see Experiment 2 Methods).

Verbal recall

After the learning task in Experiment 1, participants described each scene image from memory. To determine if there were any effects of image category on the length of verbal descriptions, we compared the mean number of words for each category using a repeated measures ANOVA with factor category. Additionally, we tested for differences between overlap and nonoverlap groups using a Welch's two-sample t-test.

Next, using the same method as for the annotation task descriptions, we defined features in verbal recall descriptions by transforming descriptions into text embeddings using MPNet, for each image and each participant. To visualize the organization of memory content in high dimensional space, we plotted the relative position of text embeddings using multidimensional scaling (MDS) based on Euclidean distance using the function dist() from the stats package in R, separately for the overlap and nonoverlap groups. To visualize the category structure, text embedding points were colored by the category of the image they were meant to describe.

Overlap-driven differences in verbal recall

Prior work assessing attraction and repulsion effects compared memory for specific features (e.g., color) to the "true" value. This provided a simple measure of how far – and in what direction – memory was biased. In this study, we use free verbal recall as our measure of memory content. However, it is more difficult to define a "ground truth" for verbal recall. As such, we devised an analysis where we assessed differences in a given participant's recall by comparing their recall to other participants' verbal descriptions. By comparing recall to a large pool of other participants, we considered other participants as a baseline sample for comparison. We applied this analysis in two ways, using either (1) nonoverlap participants as the baseline, or (2) annotation task participants as the baseline. We ran both versions to ensure any observed effects were not due to the nature of the tasks in which descriptions were obtained (i.e., memory task baseline from the nonoverlap participants versus perception task baseline from the annotation participants).

In the first version of the analysis, we computed cosine similarity between each text embedding from each participant (e.g., pool1 from sub-001) and all other text embeddings from all nonoverlap group participants. Critically, if the initial text embedding (e.g., pool1 from sub-001) was from a nonoverlap participant, we ensured similarity was only computed with text embeddings obtained from *different* nonoverlap participants (i.e., text embeddings from the remaining 59 participants in the nonoverlap group). Further, we excluded similarity values between text embeddings reflecting the same image (e.g., pool1 from sub-001 and pool1 from sub-002). Next, similarity values were Fisher Z transformed and separated into within-category (e.g., pool1 and pool2) and across-category (e.g., pool1 and library1) values. We then averaged similarity values such that we had one within-category value and one across-category value for each participant. It is important to note that some similarity values would be included in the average for more than one participant – for example, a within-category similarity value computed from two nonoverlap participants would be included in the average within-category similarity for both participants. Finally, we reverse Z transformed average similarity values and applied Welch's two-sample t-tests to determine if there was a difference in similarity values between the overlap and nonoverlap participants.

In the second version of the analysis, we performed an almost identical analysis. We first computed cosine similarity between each text embedding from each participant and the averaged text embeddings from the annotation task. Critically, we excluded similarity values between text embeddings reflecting the same image. Next, similarity values were Fisher Z transformed and separated into within-category and across-category values. We then averaged similarity values such that we had one within-category and one across-category value for each participant. Finally, we reverse Z transformed values and applied Welch's two-sample t-tests to determine if there was a difference in similarity values between the overlap and nonoverlap groups.

Results

Successful encoding across learning task

Participants' vividness ratings significantly increased across the retrieval phases of the six learning rounds (mixed effects model: $t_{(599)} = 17.50$, $p < 0.001$), approaching ceiling (i.e., a rescaled value of 1.0) by the final round ($M = 0.95$, $SD = 0.15$). While there were lower vividness ratings in the first round of learning in the overlap group compared to the nonoverlap group (Welch's

$t_{(114.83)} = 3.84$, $p < 0.001$), there was no difference between groups in the final round (Welch's $t_{(80.00)} = 0.81$, $p = 0.42$), indicating participants in the overlap group learned image associations just as well as participants in the nonoverlap group. Accuracy on the multiple-choice test was an inclusion criterion for Experiment 1 (see Methods). Taken together, these results suggest participants were able to successfully learn the face-scene associates.

Verbal descriptions

In the annotation task, participants described scene images from six different semantic categories. The mean number of words produced was 24.77 ($SD = 7.35$). We compared the mean number of words in verbal descriptions for each category and found no difference between categories in a repeated measures ANOVA ($F_{(5,1344)} = 0.48$, $p = 0.79$), indicating category did not affect this baseline measure.

Verbal recall

After learning in Experiment 1, participants described scene images from memory in a verbal recall task. The mean number of words produced was 30.78 ($SD = 15.65$). Furthermore, a two-sample t-test revealed no difference between groups (Welch's $t_{(117.72)} = 0.015$, $p = 0.99$). We next compared the mean number of words in verbal recall for each category and found no difference between categories in a repeated measures ANOVA ($F_{(5,414)} = 1.47$, $p = 0.20$). Finally, we transformed verbal descriptions into text embeddings (Fig. 2.2a) and visualized the organization of recall descriptions across the six categories in each group (Fig. 2.2b shows category structure in an MDS plot).

Overlap-driven differences in verbal recall

In Experiment 1, we sought to compare verbal recall in the overlap and nonoverlap groups. The critical test was whether recall similarity within a category was affected by group (i.e., the presence of overlapping stimuli). Higher within-category similarity in the overlap group would indicate verbal recall becomes more similar in response to overlap; lower within-category similarity in the overlap group would indicate verbal recall becomes more separated in response to overlap. We performed this analysis using two different baseline measures when computing within-category (and across-category) similarity values.

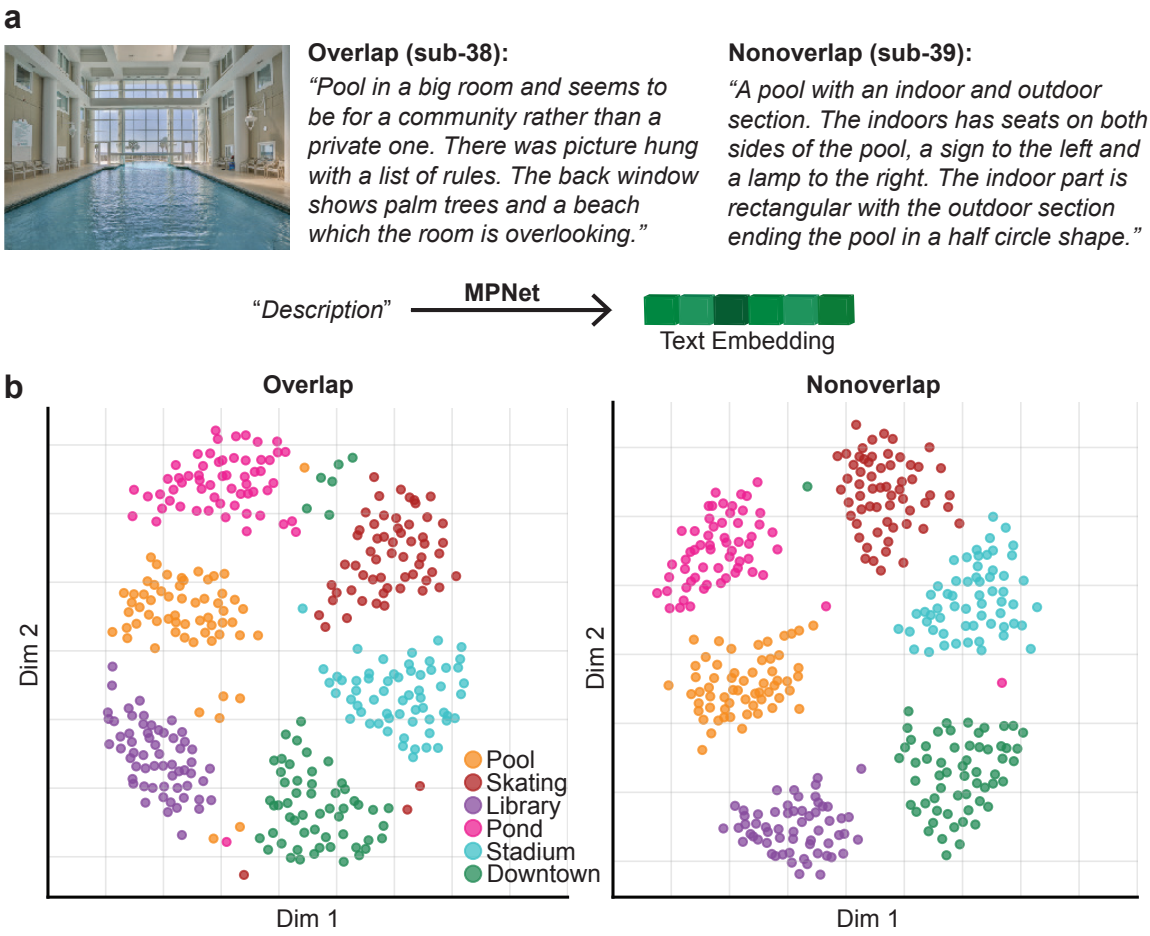

Figure 2.2. Organization of recall in Experiment 1. a) Verbal recall. Sample verbal recall descriptions from participants in the overlap group and the nonoverlap group. Recall descriptions were transformed into text embeddings using MPNet. **b)** MDS plots of text embeddings from participants in the overlap group (left) and nonoverlap group (right), colored by image category.

In the first version of this analysis, we used other nonoverlap participants' recall as our baseline for comparison. In a two-sample t-test, we found that within-category similarity was lower in the overlap group (0.65 ± 0.07) compared to the nonoverlap group (0.69 ± 0.02; Welch's $t_{(72.30)} = 4.08$, $p < 0.001$; Fig. 2.3, left). Meanwhile, across-category similarity was higher in the overlap group (0.23 ± 0.04) compared to the nonoverlap group (0.21 ± 0.02; Welch's $t_{(76.20)} = 2.24$, $p = 0.028$). In the second version of this analysis, we used annotation task descriptions as our baseline for comparison. As in the first version, we found that within-category similarity was lower in the overlap group (0.74 ± 0.07) compared to the nonoverlap group (0.79 ± 0.03; Welch's $t_{(81.04)} = 5.18$, $p < 0.001$; Fig. 2.3, right). Meanwhile, there was no difference in across-category similarity between the overlap group (0.24 ± 0.02) and nonoverlap group (0.25 ± 0.05; Welch's $t_{(76.17)} = 1.67$,

p = 0.10). As such, we did not place much weight on the significantly higher across-category similarity in the overlap group when using the nonoverlap baseline, as this result was not consistent across the two baseline measures.

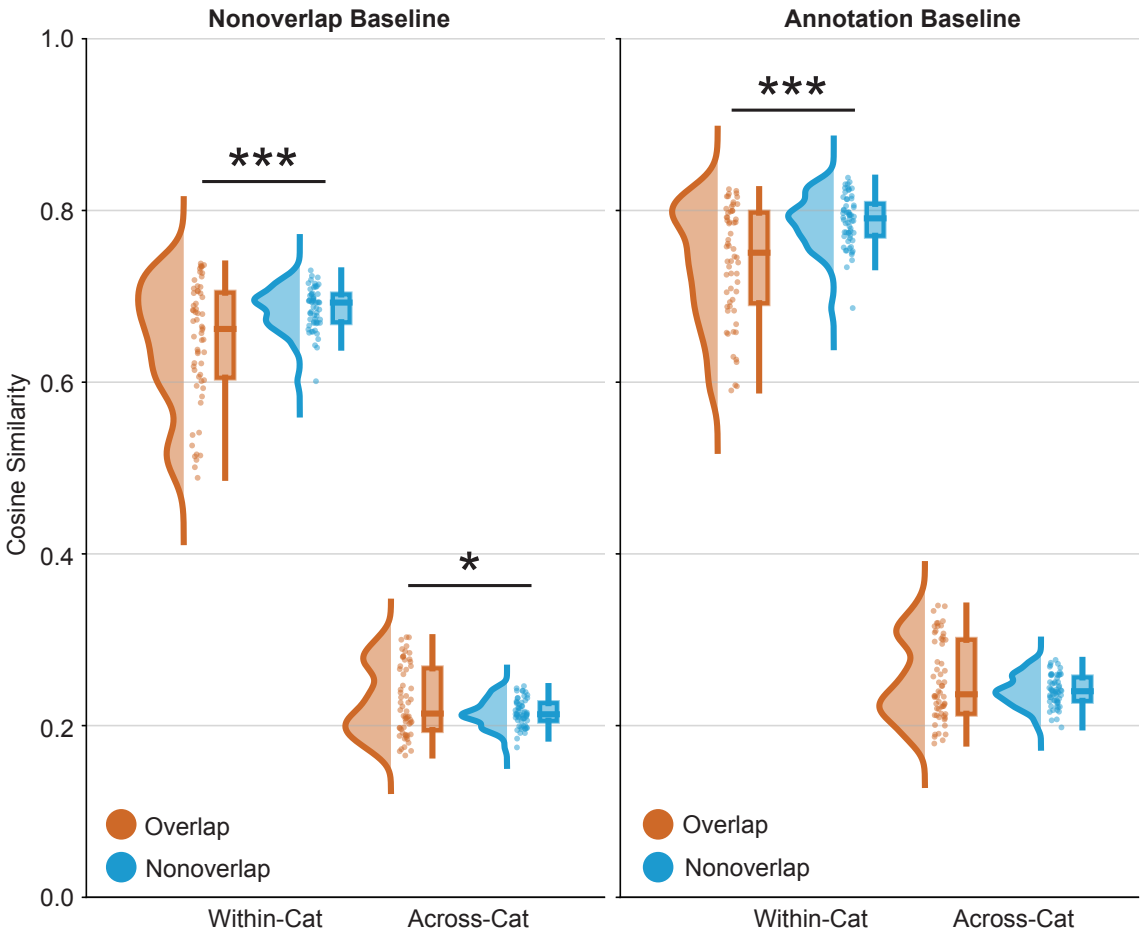

Figure 2.3. Experiment 1 results. Cosine similarity between within-category and across-category images in the overlap group (orange) and nonoverlap group (blue). Left panel depicts results using nonoverlap group as the baseline and right panel depicts results using the annotation task as the baseline.

Taken together, these results indicate overlap triggers a separation in memory content wherein overlapping scene images from the same category were described more differently from one another. However, since there are multiple competing stimuli, each possessing a different subset of overlapping and nonoverlapping features, it is impossible to determine whether differences in verbal recall reflect targeted separation from a specific competitor versus a general process in response to overlap (e.g., increase in memory detail). Moreover, with multiple

overlapping competitors, targeted separation may not be feasible, as movement *away* from one overlapping image may necessitate movement *toward* another. Therefore, we designed a second experiment with only two competitors in each category.

Experiment 2

Experiment 1 indicated overlap leads to separation in verbal recall. This finding is analogous to prior findings of repulsion, where memories for overlapping stimuli are adaptively separated to prevent memory interference. However, since Experiment 1 included multiple overlapping stimuli within a category, it is impossible to determine whether the separation in verbal recall reflects targeted separation versus a general response to overlap. Therefore, in Experiment 2, we included only two overlapping scene images. We hypothesized that if targeted separation is indeed occurring in verbal recall, we would observe differences in verbal recall that are specific to a particular studied competitor image. Meanwhile, if verbal recall is affected by a more general mechanism, we would expect no difference in verbal recall based on the specific studied competitor.

Methods

In Experiment 2, participants each studied pairs of images from six different semantic categories. We used data from the annotation task (see Experiment 1 Methods) in order to select images for Experiment 2. Critically, one of the images in each category was chosen as the most 'central' image within a given category and was common across all participants (i.e., target image). The other image from each category was assigned as one of the remaining five possible images in a given category (i.e., competitor image).

Participants

For Experiment 2, target sample size was identified in advance, and additional participants were recruited to reach the target after applying exclusion criteria. To reach these samples, participants were recruited from the University of Oregon subject pool. All participants self-reported to be right-handed, native English speakers, and have normal or corrected-to-normal vision. Informed consent was obtained from all participants according to a protocol approved by the University of Oregon Institutional Review Board, and all participants received credit for

participation. Because we did not have a way to estimate effect sizes in advance, we chose a target sample of 90 participants (60 female; mean age 19.07 ± 1.23 SD, range 18-35). Exclusion criteria are described in the Experimental Procedure section.

Stimuli

Thirty-six scene images and twelve face images were used in Experiment 2. The 36 scene images were selected equally from the six categories in the annotation task such that each category included six images. Within each category, there was one "target" image common across all participants and five possible "competitor" images. Since each participant studied one target and one competitor per category, each participant was assigned 12 images in total. Each of the twelve scene images was randomly paired with a face image, creating a set of twelve face-scene paired associates for each subject. The faces served as cues during cued recall tasks (described below).

By manipulating which competitor a given participant was assigned, we hoped to manipulate *how* recall for the target image would be influenced (i.e., relative to the actual competitor studied). As such we had two primary goals when selecting stimuli: (1) inducing competition from highly overlapping stimuli, and (2) selecting a target image that was "central" to the possible competitor images in text embedding space. To accomplish the first goal, we computed pairwise correlations between averaged annotation task text embeddings for all possible six-image subsets (out of the 15 total images within a category) and selected the subset of images that were most similar to one another. To accomplish the second goal, we determined which of the six selected images was most correlated with the other five images (i.e., the most "central" image). The image with the highest mean correlation with the other five images was chosen to be the common target image and the remaining five images were chosen to be the competitor images. To visualize this selection process, we plotted the relative position of all averaged text embeddings in the annotation task in two-dimensional space using MDS based on Euclidean distance (Fig. 2.4a). To visualize the category structure, text embedding points were colored by the category of the image they are meant to describe. Next, we plotted the relative position of selected text embeddings from a single category (stadium) using MDS (Fig. 2.4b).

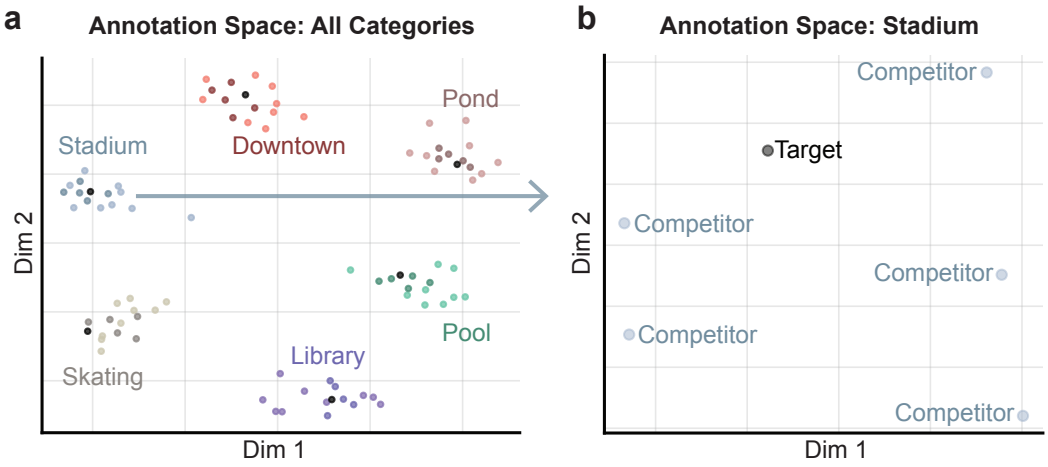

Figure 2.4. Organization of descriptions in annotation task. a) MDS plot of averaged text embeddings from participants in the annotation task, colored by category. Black points in each category reflect the selected target image and darker colored points reflect the five competitor images. **b)** MDS plot of selected target and competitor text embeddings from stadium category.

Experimental procedure

In Experiment 2, participants were assigned 12 scene images with two images drawn from each category. Critically, one of these images (i.e., "target") was common across all participants, and the other image (i.e., "competitor") was counterbalanced across participants from the remaining five images in a given category (Fig. 2.5a). As such, participants were evenly distributed across the five possible competitor images (i.e., 18 participants per competitor image).

Participants completed six runs of a learning task (Fig. 2.5b). Each run began with a self-initiated start screen and included three phases: encoding, retrieval, and test. The additional test phase in Experiment 2 was added in order to facilitate learning for twice the number of associations required in Experiment 1. The start of each phase was preceded by a screen introducing the phase as "Study," "Vividness," or "Test" (i.e., encoding, retrieval, and test; 5500ms). Across all phases, stimuli were displayed on a white background. After completing the learning task, participants completed two additional memory tasks. The experiment was implemented in PsychoPy and jsPsych and hosted on Pavlovia (https://pavlovia.org), and lasted ~1h30min.

The encoding and retrieval phases were identical to those of Experiment 1. During encoding phases, participants learned all face-scene associations, one association at a time. During retrieval phases, participants were cued to recall the scene associated with each of the faces, one association at a time. During the test phases, participants completed a multiple-choice recognition

test where they were presented with each scene image at the top of the screen and four face images below. One face image lure was paired with the other same-category scene image, and the two remaining face image lures were paired with the two scene images from a randomly selected category. Participants were asked to click on the correct face image using the mouse. The test phase was separated into two blocks, with each scene presented once within each block. The order in which the face images were tested was randomized within each run and block, and for each participant. Due to a technical error, behavioral data from the retrieval and test phases were not saved and therefore not considered in analyses.

After completing the learning task, participants completed a cued recall test identical to that of Experiment 1. For each face-scene associate, participants were shown a cue (face) at the top of the screen and were asked to type a description of the associated scene image (Fig. 2.5c). Following cued verbal recall, participants completed a multiple-choice recognition test for each of the face-scene associates that was nearly identical to that of Experiment 1. In Experiment 2, the six scene images comprised of two images from each of three categories: the target category and two randomly selected categories. Participants missing any verbal recall response and/or incorrectly responding to more than one multiple-choice question were excluded from analysis (n=6). Additional participants were recruited to make up for these participants to maintain balanced participants for each of the five possible competitors.

Verbal recall

After the learning task in Experiment 2, participants described each scene image from memory. To determine if there were any effects of image category on verbal descriptions, we compared the mean number of words for each category using a repeated measures ANOVA with factor category. As in Experiment 1, we defined features in verbal descriptions by transforming descriptions into text embeddings using MPNet, yielding 768-dimensional vectors representing semantic information in verbal descriptions for each image and each participant.

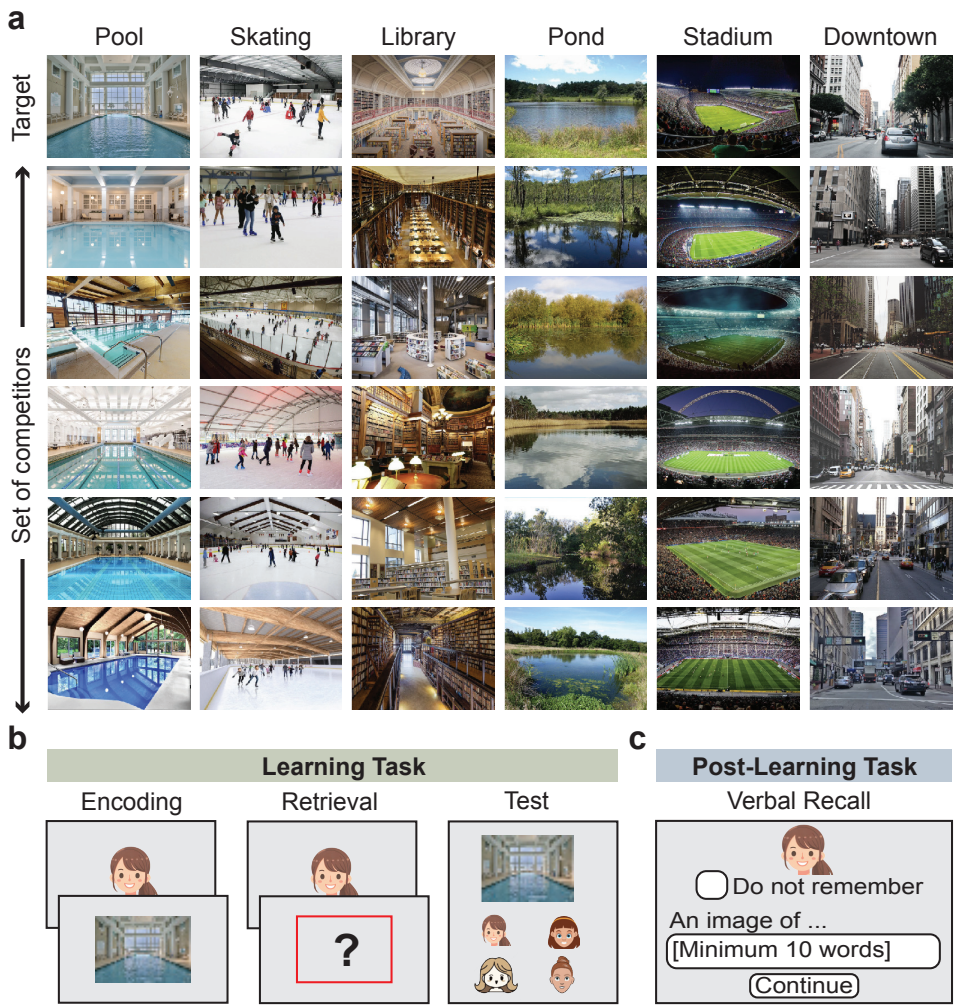

Figure 2.5. Stimuli & task design for Experiment 2. a) Stimuli. Scene images used in Experiment 2. Participants were assigned the target image as well as one of five possible competitor images in each category. **b)** Learning task. Learning consisted of encoding, retrieval, and test trials. **c)** Recall task. After learning, participants were cued with each face image at the top of the screen and asked to type a description of the associated scene image.

Specific overlap-driven differences in verbal recall

As with Experiment 1, we assessed differences in a given participant's recall by comparing their recall to other participants' verbal descriptions. To assess whether there was a targeted effect of the specific studied competitor on recall of the target images, we compared similarity in verbal recall across participants that studied (1) the same competitor, or (2) other possible competitors. First, we compared similarity between text embeddings for target recall and text embeddings for competitor recall. As before, we computed cosine similarity between each target text embedding and all other participants' competitor text embeddings. Critically, we separated similarity values

based on whether the other participant had been assigned the same competitor or another competitor. Cosine similarity values were Fisher Z transformed and averaged for each participant such that each participant had two target-competitor similarity values: one for the same competitor and one for other possible competitors. Next, we reverse Z transformed values and ran a paired sample t-test to determine whether verbal recall for the target was more similar to versus more different from the specific competitor studied.

We next considered whether the specific competitor could be influencing target memory in a consistent but nondirectional manner. To do so, we compared consistency for target memory based on its competitor. We first computed cosine similarity between each target text embedding and all other participants' target text embeddings. Critically, we separated similarity values based on whether the other participant had been assigned the same competitor or another competitor. Cosine similarity values were Fisher Z transformed and averaged for each participant such that each participant possessed two target-target similarity values (one for same competitor and one for other competitors). Next, we reverse Z transformed values and ran a paired sample t-test to determine whether target recall was more consistent among participants sharing the same competitor than among participants with other competitors.

Results

Successful encoding across learning task

Accuracy on the multiple-choice test was an inclusion criterion for Experiment 2 (no more than one error). Participants incorrectly responding to more than one multiple-choice question and/or missing any verbal recall response were excluded from analysis (n=6).

Image descriptions as a measure of memory content

After learning in Experiment 2, participants described scene images from memory in a verbal recall task. The mean number of words produced was 25.81 (SD = 14.52). The mean number of words produced during verbal recall did not differ across categories (repeated measures ANOVA, $F_{(5,534)} = 0.72$, $p = 0.61$).

Specific overlap-driven differences in verbal recall

In Experiment 2, participants studied and recalled a common target image and a variable competitor image from each category. We first considered whether memory for the target image was specifically drawn toward versus pushed away from the competitor it was paired with. To do so, we compared similarity between recall of each target with recall of the same competitor it was paired with versus other possible competitors. We found no difference in target-competitor similarity when the competitor was the same competitor studied versus other competitors in a paired-samples t-test ($t_{(89)} = 0.81$, $p = 0.42$; Fig. 2.6, left). This finding suggests memory for the target image is neither drawn toward nor pushed away from its specific overlapping competitor image, relative to other possible competitors.

We next considered whether the specific competitor could be influencing target memory in a different manner. To do so, we compared consistency for target memory based on the competitor it was paired with. If there was no difference in target-target similarity for targets studied alongside the same versus other competitors, that would indicate the specific competing image did not affect memory in a consistent way. However, if there were higher target-target similarity values when participants studied the same competitor, this would indicate the specific competitor affected target memory in a consistent (but not necessarily directional) manner. Lower target-target similarity values for the same competitor would be difficult to interpret. We found that similarity was significantly higher when both targets had been studied alongside the same competitor than another competitor in a paired-samples t-test ($t_{(89)} = 4.26$, $p < 0.001$; Fig. 2.6, right). This finding suggests that there exists some mechanism by which the specific competitor systematically alters memory for the target in a consistent manner. A possible interpretation of this effect is that the specific competitor affects target recall in different ways for different features. Specifically, there may be a combination of attraction and separation effects where some shared features and some diagnostic features are simultaneously prioritized, leading to a consistent but net nondirectional effect.

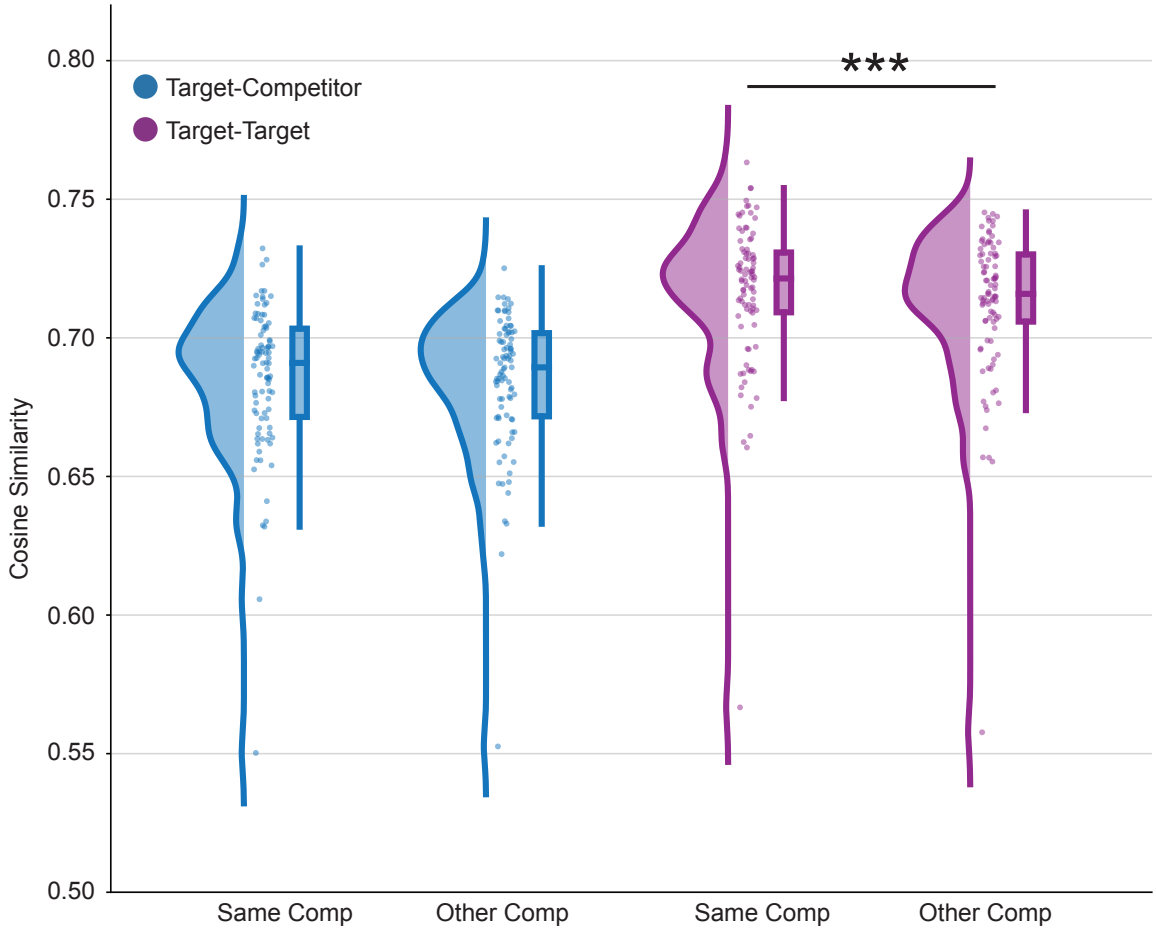

Figure 2.6. Experiment 2 results. Cosine similarity between target-competitor (left) and target-target (right) text embeddings based on whether target was paired with the same competitor or other possible competitor.

General Discussion

In this study, we conducted a pair of behavioral experiments assessing the influence of overlap on verbal recall. In both experiments, we used naturalistic scene stimuli and collected verbal recall as our measure of memory content. In Experiment 1, we showed that overlap from multiple competing stimuli leads to separation in verbal recall. However, it remains unclear whether this separation was targeted based on the specific studied images versus a general response to overlap. As such, we conducted Experiment 2 where we considered how verbal recall for a single target image differs based on the specific overlapping competitor it was paired with. Results from Experiment 2 suggest verbal recall for a given image is indeed impacted by the specific competitor – but not in a specific direction. In summary, these findings indicate overlap can lead

to (1) separated recall for multiple overlapping images, and (2) targeted effects based on a specific overlapping competitor. Taken together, these findings reveal subtle and variable influences of overlap on verbal recall for naturalistic images.

A key contrast between the two experiments is that we observe a separation in verbal recall for overlapping images in Experiment 1 (analogous with behavioral repulsion), but no evidence for separation in Experiment 2. We present two interpretations for this discrepancy that reflect methodological limitations. One interpretation is that separation only occurred in Experiment 1 due to the relatively greater difficulty of the task (learning six overlapping images in Experiment 1 versus only two overlapping images in Experiment 2). This possibility is consistent with Chanales et al. (2021), which shows that repulsion occurs only when discrimination between overlapping stimuli is sufficiently difficult. Another interpretation is that separation occurred in both experiments, but the effects were too subtle to be detected in Experiment 2. Specifically, since Experiment 1 compared participants that experienced overlap to participants that experienced no overlap and Experiment 2 compared participants that experienced overlap to one another, it would be more difficult to detect the subtler differences in Experiment 2. Moreover, the comparatively more subtle effect in Experiment 2 may consist of simultaneous attraction and separation processes that effectively cancel one another other out. As such, comparing target recall directly to competitor recall may be an inadequate method for assessing competitor-driven effects.

An alternative theoretical interpretation of our results is that the repulsion-like effects in Experiment 1 do not actually reflect repulsion. Instead, it may be that overlap triggered a more general 'sharpening' of memory wherein participants recalled more item-specific details for the overlapping scenes. An increase in item-specific memory would manifest as the separation of recall observed in Experiment 1 and the lack of effect observed in Experiment 2. Supporting this idea, Drascher & Kuhl (2022) showed that overlap can not only lead to repulsion in memory, but also increased precision in memory for facial features – analogous to an increase in item-specific memory in verbal recall. Neuroimaging studies further support this interpretation by demonstrating that increased representational consistency (i.e., stronger item-specific representations) during encoding is predictive of improved learning (Lu et al., 2015; Ward et al., 2013; Xue et al., 2010, 2013). As such, it may be that overlap triggers an adaptive increase in item-specific representations – in both the brain and behavior – in order to overcome competition in memory.

The current study was largely inspired by prior studies demonstrating behavioral repulsion of memory content (Chanales et al., 2021; Drascher & Kuhl, 2022). These studies suggest repulsion of memory content may be a downstream effect of differentiation of hippocampal representations (Chanales et al., 2017; Dimsdale-Zucker et al., 2018; Favila et al., 2016; Hulbert & Norman, 2015; Kim et al., 2017; Schlichting et al., 2015; Wanjia et al., 2021), which also occurs in response to overlapping stimuli. Yet in the current study, we fail to find definitive evidence of behavioral repulsion. This discrepancy raises the question of whether our findings are actually consistent with neuroimaging studies. One explanation is that hippocampal differentiation occurs dynamically depending on the organization and complexity of the stimuli. A recent study from Mirjalili et al. (2025) indicates that hippocampal differentiation for multiple overlapping competitors does not happen simultaneously. Instead, stimuli tend to be separated along different diagnostic dimensions at different stages in learning. Therefore, if hippocampal differentiation is indeed predictive of repulsion in memory content, it may be that we should not expect targeted separation of all competitors along all features at the moment of recall. Drascher & Kuhl (2022) exemplify this possibility by demonstrating repulsion in memory content specifically occurs along a diagnostic (i.e., discriminating) feature. Verbal recall, with its hundreds of feature dimensions, may be obscuring more subtle biases along specific features.

The use of verbal recall as our measure of memory content introduces multiple sources of variability. As such, it is difficult to draw direct comparisons between the present work and prior studies. For one, prior studies of attraction and repulsion probed memory along one or two specific features (e.g., color; Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024). Verbal recall, quantified via large language models, possesses hundreds of feature dimensions, which increases the dimensionality and noise in the data. As such, subtle differences in verbal recall must be particularly strong in order to be detected. Additionally, verbal recall differs from prior studies in that the features captured by large language models are *implicitly* measured. For example, one feature in a text embedding may be the extent of 'manmade-ness' in the description. When participants are prompted to describe a given scene image, they are not explicitly tasked with describing how manmade the scene is. In contrast, prior studies probing memory for specific features explicitly ask participants to report the color (or another feature) of a stimulus (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024). These features were typically the same features along which overlapping stimuli could

be discriminated. By using variable and naturalistic stimuli in our study, it would be impossible to explicitly probe memory for every single feature, let alone identify diagnostic features for each image. Therefore, an explicit measure such as those used in prior studies would be insufficient to quantify variability in our stimulus set. Instead, our study provides insight into how verbal recall – specifically – is influenced by overlapping experiences.

A unique finding in Experiment 2 was a targeted effect of specific competitor that reflects neither attraction nor repulsion. Specifically, we found that recall for the target image was consistent based on the overlapping competitor it was paired with, but not specifically towards or away from that competitor. As such, Experiment 2 indicates overlap can influence memory content in ways beyond what has been established in prior studies. While we cannot form strong conclusions from these data, we speculate on a possible mechanism. The scene stimuli used in this study possess a large amount of variability, with different sources of salience – for example, a competitor image from the downtown street category may include a bright red car in the bottom right corner. This salient feature could direct attention to the same spatial location (bottom right) in the target image, regardless of whether there was a salient feature in that location. In this manner, recall of the target could be consistently influenced based on its specific competitor by driving attention to different features (Fan & Turk-Browne, 2013; Uncapher & Rugg, 2009). Moreover, we could conceptualize this effect as spatially driven – where verbal recall prioritizes specific parts of an image with salient features. In other words, the salient and diagnostic features in a competitor image may direct attention during encoding (and therefore recall) of the target image. To assess this possibility, we develop a new method for assessing competitor-specific differences in recall in Chapter III. Through this work, we may tease apart the multifaceted ways in which overlap alters memory content.

In summary, we provide evidence for separation in verbal recall due to overlap in Experiment 1 and specific competitor-driven verbal recall in Experiment 2. In both experiments, we use verbal recall as a rich and complex measure of memory for naturalistic scene images. While verbal recall is a useful tool for quantifying more naturalistic memories, it possesses its own unique organization and variability, which can affect the nature of observable memory effects. As such, the present work serves to further our understanding of naturalistic memories within the context of verbal recall. With different measures of behavioral memory, it is possible that we would find much different effects – future work could compare these different measures to better understand

the dynamic influences of overlap on memory. Moreover, we have yet to directly link fMRI measures of neural activity to verbal recall of overlapping memories, which could reveal whether neural activity patterns during encoding are sensitive to diagnostic features in subsequent verbal recall. In conclusion, adaptive changes in memory are flexible and variable, depending on both the nature of stimuli and the way in which memory is probed. Through the experiments in this chapter, we demonstrate that subtle effects using naturalistic stimuli can be quantified using verbal recall, providing insight into how overlap could affect naturalistic memories more broadly.

References Cited

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*(1), 68–100. https://doi.org/10.1037/0033-295X.102.1.68

Antony, J. W., Romero, A., Vierra, A. H., Luenser, R. S., Hawkins, R. D., & Bennion, K. A. (2022). Semantic relatedness retroactively boosts memory and promotes memory interdependence across episodes. *eLife*, *11*, e72519. https://doi.org/10.7554/eLife.72519

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

Chanales, A. J. H., Oza, A., Favila, S. E., & Kuhl, B. A. (2017). Overlap among Spatial Memories Triggers Repulsion of Hippocampal Representations. *Current Biology*, *27*(15), 2307-2317.e5. https://doi.org/10.1016/j.cub.2017.06.057

Chanales, A. J. H., Tremblay-McGaw, A. G., Drascher, M. L., & Kuhl, B. A. (2021). Adaptive Repulsion of Long-Term Memory Representations Is Triggered by Event Similarity. *Psychological Science*, *32*(5), 705–720. https://doi.org/10.1177/0956797620972490

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Cohn-Sheehy, B. I., Delarazan, A. I., Reagh, Z. M., Crivelli-Decker, J. E., Kim, K., Barnett, A. J., Zacks, J. M., & Ranganath, C. (2021). The hippocampus constructs narrative memories across distant events. *Current Biology: CB*, *31*(22), 4935-4945.e7. https://doi.org/10.1016/j.cub.2021.09.013

Dimsdale-Zucker, H. R., Ritchey, M., Ekstrom, A. D., Yonelinas, A. P., & Ranganath, C. (2018). CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nature Communications*, *9*(1), 294. https://doi.org/10.1038/s41467-017-02752-1

Drascher, M. L., & Kuhl, B. A. (2022). Long-term memory interference is resolved via repulsion and precision along diagnostic memory dimensions. *Psychonomic Bulletin & Review*, *29*(5), 1898–1912. https://doi.org/10.3758/s13423-022-02082-4

Fan, J. E., & Turk-Browne, N. B. (2013). Internal attention to features in visual short-term memory guides object learning. *Cognition*, *129*(2), 292–308. https://doi.org/10.1016/j.cognition.2013.06.007

Favila, S. E., Chanales, A. J. H., & Kuhl, B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nature Communications*, *7*(1), 11066. https://doi.org/10.1038/ncomms11066

Hebscher, M., Bainbridge, W. A., & Voss, J. L. (2023). Neural similarity between overlapping events at learning differentially affects reinstatement across the cortex. *NeuroImage*, *277*, 120220. https://doi.org/10.1016/j.neuroimage.2023.120220

Hulbert, J. C., & Norman, K. A. (2015). Neural Differentiation Tracks Improved Recall of Competing Memories Following Interleaved Study and Retrieval Practice. *Cerebral Cortex*, *25*(10), 3994–4008. https://doi.org/10.1093/cercor/bhu284

Kim, G., Norman, K. A., & Turk-Browne, N. B. (2017). Neural Differentiation of Incorrectly Predicted Memories. *The Journal of Neuroscience*, *37*(8), 2022–2031. https://doi.org/10.1523/JNEUROSCI.3272-16.2017

Leeuw, J. R. de, Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an Open-Source Collaborative Ecosystem of Behavioral Experiments. *Journal of Open Source Software*, *8*(85), 5351. https://doi.org/10.21105/joss.05351

Lu, Y., Wang, C., Chen, C., & Xue, G. (2015). Spatiotemporal Neural Pattern Similarity Supports Episodic Memory. *Current Biology*, *25*(6), 780–785. https://doi.org/10.1016/j.cub.2015.01.055

Mensink, G.-J., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review*, *95*(4), 434–455. https://doi.org/10.1037/0033-295X.95.4.434

Mirjalili, S., Wanjia, G., Grätz, D., Wang, E., Mayr, U., & Kuhl, B. A. (2025). Hippocampal transformations occur along dimensions of memory interference. *bioRxiv*, 2025.10.13.682242. https://doi.org/10.1101/2025.10.13.682242

Morales-Torres, R., Wing, E. A., Deng, L., Davis, S. W., & Cabeza, R. (2024). Visual Recognition Memory of Scenes Is Driven by Categorical, Not Sensory, Visual Representations. *Journal of Neuroscience*, *44*(21). https://doi.org/10.1523/JNEUROSCI.1479-23.2024

Morrow, E., & Clewett, D. (2025). Distortion of overlapping memories relates to arousal and anxiety. *Cognitive, Affective & Behavioral Neuroscience*, *25*(1), 154–172. https://doi.org/10.3758/s13415-024-01229-1

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review*, *56*(3), 132–143. https://doi.org/10.1037/h0057488

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Reagh, Z. M., & Ranganath, C. (2023). Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. *Nature Communications*, *14*(1), 1279. https://doi.org/10.1038/s41467-023-36805-5

Rehurek, R. (2011). Gensim–python framework for vector space modelling. In *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic.* (Vol. 3, Issue 2). https://cir.nii.ac.jp/crid/1370869855565691916

Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks* (arXiv:1908.10084). arXiv. https://doi.org/10.48550/arXiv.1908.10084

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*(1), 8151. https://doi.org/10.1038/ncomms9151

Shohamy, D., & Wagner, A. D. (2008). Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. *Neuron*, *60*(2), 378–389. https://doi.org/10.1016/j.neuron.2008.09.023

Song, K., Tan, X., Qin, T., Lu, J., & Liu, T.-Y. (2020). *MPNet: Masked and Permuted Pre-training for Language Understanding* (arXiv:2004.09297). arXiv. https://doi.org/10.48550/arXiv.2004.09297

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event Congruency Enhances Episodic Memory Encoding through Semantic Elaboration and Relational Binding. *Cerebral Cortex (New York, NY)*, *19*(5), 1198–1207. https://doi.org/10.1093/cercor/bhn165

Tandoc, M. C., Dong, C. V., & Schapiro, A. C. (2024). Object Feature Memory Is Distorted by Category Structure. *Open Mind: Discoveries in Cognitive Science*, *8*, 1348–1368. https://doi.org/10.1162/opmi_a_00170

Uncapher, M. R., & Rugg, M. D. (2009). Selecting for Memory? The Influence of Selective Attention on the Mnemonic Binding of Contextual Information. *The Journal of Neuroscience*, *29*(25), 8270–8279. https://doi.org/10.1523/JNEUROSCI.1043-09.2009

Wang, Y., Lee, H., & Kuhl, B. A. (2023). *Mapping multidimensional content representations to neural and behavioral expressions of episodic memory* [Preprint]. Neuroscience. https://doi.org/10.1101/2023.01.09.523096

Wanjia, G., Favila, S. E., Kim, G., Molitor, R. J., & Kuhl, B. A. (2021). Abrupt hippocampal remapping signals resolution of memory interference. *Nature Communications*, *12*(1), 4816. https://doi.org/10.1038/s41467-021-25126-0

Ward, E. J., Chun, M. M., & Kuhl, B. A. (2013). Repetition Suppression and Multi-Voxel Pattern Similarity Differentially Track Implicit and Explicit Visual Memory. *Journal of Neuroscience*, *33*(37), 14749–14757. https://doi.org/10.1523/JNEUROSCI.4889-12.2013

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

Xue, G., Dong, Q., Chen, C., Lu, Z.-L., Mumford, J. A., & Poldrack, R. A. (2013). Complementary Role of Frontoparietal Activity and Cortical Pattern Similarity in Successful Episodic Memory Encoding. *Cerebral Cortex*, *23*(7), 1562–1571. https://doi.org/10.1093/cercor/bhs143

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and Ventral Medial Prefrontal Activation during Retrieval-Mediated Learning Supports Novel Inference. *Neuron*, *75*(1), 168–179. https://doi.org/10.1016/j.neuron.2012.05.010

Zhao, Y., Chanales, A. J. H., & Kuhl, B. A. (2021). Adaptive Memory Distortions Are Predicted by Feature Representations in Parietal Cortex. *The Journal of Neuroscience*, *41*(13), 3014–3024. https://doi.org/10.1523/JNEUROSCI.2875-20.2021

CHAPTER III
VERBAL RECALL OF OVERLAPPING SCENES PRIORITIZES DISSIMILAR IMAGE PATCHES

Introduction

Extensive work has demonstrated that overlap can affect the way visual stimuli are remembered relative to one another (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024). However, visual memory tasks have been largely constrained to artificially generated stimuli and retrieval of simple features (e.g., color or orientation; Bae & Luck, 2017; Chanales et al., 2021; Chapman et al., 2023; Chunharas et al., 2022; Golomb, 2015; Morrow & Clewett, 2025; Rademaker et al., 2015). As a result, memory for multiple features must be probed via multiple retrieval tasks. In Chapter II, on the other hand, we used verbal recall descriptions (transformed into text embeddings) as our measure of memory content. Verbal recall is unconstrained, allowing participants to recall as many features as they choose within a single retrieval task. However, the broader feature space of text embeddings affords its own limitation: text embeddings are not particularly interpretable, making it difficult to characterize what features a participant is actually remembering.

Using verbal recall text embeddings in Chapter II, we observed some results that were difficult to reconcile. Specifically, Experiment 1 revealed that verbal recall from participants experiencing overlap was more separated compared to participants experiencing no overlap. Meanwhile, Experiment 2 found that verbal recall was consistent based on the specific competitor image it was paired with but reflected neither attraction to nor separation from that competitor. One possible explanation for this discrepancy is that effects were more difficult to capture in Experiment 2, which compared participants that experienced different types of overlap to one another. As such, it may be that verbal recall on its own was not sensitive enough to detect the relatively more subtle effects in Experiment 2. To better understand these findings, the current chapter develops a new method to translate verbal recall into an interpretable feature space that can reveal the nature of subtle differences in recall.

A key benefit of the constrained artificial stimuli used in prior studies was that their diagnostic features (e.g., color) could be explicitly controlled (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024). As such, memory biases manifested along those same controlled features. With naturalistic scene stimuli, on the other hand, it is

difficult to identify all diagnostic features, let alone control for them. Some studies attempt to quantify features of naturalistic scene images by probing memory for objects within a scene (Hollingworth, 2005) and/or spatial layout of drawings (Bainbridge et al., 2019). However, verbal recall (as used in Chapter II) does not easily translate to the 2D space of a scene image, as verbal recall inherently reflects semantic rather than visual information. As such, this chapter seeks to translate verbal recall from Experiment 2 into an interpretable 2D feature space and ultimately assess whether this new feature space reflects diagnostic parts of overlapping stimuli. Specifically, we segmented target images into smaller patches and assessed how well verbal recall maps to each of these patches. For example, if a competitor image from the downtown street category included a bright red car in the bottom right corner, an image patch in the bottom right corner of the target may be prioritized in verbal recall. In this manner, we could assess whether the importance of a given patch in verbal recall was driven by the specific competitor image that was studied.

In the current chapter, we tested whether overlapping stimuli influence verbal recall along specific image patches. To do so, we reanalyzed data from Experiment 2 of Chapter II. Specifically, we segmented target images into smaller patches and assessed the extent to which each patch contributed to the distinctiveness of recall. Critically, we compared distinctiveness of recall at different patch resolutions (i.e., larger image segments versus smaller image segments), revealing the spatial scale at which verbal recall captures relevant features (e.g., structural image details versus fine-grained localized details). Next, we assessed whether the organization of patch importance in verbal recall was driven by the specific competitor image that was studied. Finally, we considered whether the importance of different patches may be driven by 'objective' dissimilarity between the target and competitor in that patch using deep vision model VGG-16.

Methods

Experimental design

In the current experiment, N=90 participants each studied pairs of images from six different semantic categories (e.g., pools). Critically, one of the images in each category was chosen as the most 'central' image within a given category and was common across all participants (i.e., target image). The other image from each category was alternately assigned from the remaining five possible images in that category (i.e., competitor image), resulting in a set of 12 images for each participant. See Chapter II Experiment 2 Methods for full description.

Participants completed six runs of a learning task (Fig. 2.5b) where they learned to pair each scene image with a unique face image across encoding, retrieval, and test phases. During encoding phases, participants were presented with all face-scene associations, one association at a time. During retrieval phases, participants were cued to recall the scene associated with each of the faces, one association at a time. During the test phases, participants completed a multiple-choice recognition test where they were presented with each scene image at the top of the screen and four face images below and were asked to click on the correct face.

After completing the six rounds of the learning task, participants completed a cued recall test. For each face-scene associate, participants were shown the face cue at the top of the screen and were asked to type a description of the associated scene image (Fig. 2.5c). As before, we defined features in verbal descriptions by transforming descriptions into text embeddings using large language model MPNet (Reimers & Gurevych, 2019; Song et al., 2019), yielding 768-dimensional vectors representing semantic information in verbal descriptions for each image and each participant. Following cued verbal recall, participants completed a multiple-choice recognition test for each of the face-scene associates.

Image segmentation

We segmented each image (500 x 700 pixels) into N×N patch grids, ranging from N=2 to N=5 resolution (e.g., Fig. 3.1, left). We chose a range of resolutions to account for different scales of spatial organization within an image (e.g., structural layout of an indoor pool versus placement of chair beside pool). Following segmentation, 'objective' features in each patch were quantified using convolutional neural network VGG-16 (Simonyan & Zisserman, 2015). Specifically, each patch was first converted into array format, then input into VGG-16 using TensorFlow and Keras libraries (Abadi et al., 2016; Chollet, 2015). Next, we extracted features from an early, middle, and late layer of processing (block 1 pool, block 3 pool, block 5 pool). This yielded 802814-, 200702-, and 25088-dimensional vectors, respectively, representing features that vary from more visual (early layers) to more semantic (late layers).

Additionally, we generated verbal descriptions for each patch in each target image using the Qwen3-VL model (Bai et al., 2025) as a measure of semantic features. These descriptions were later used for comparison to verbal recall descriptions of the target. The Qwen3-VL model was instructed to generate one to two sentences "sounding like a human casually recalling the image

from memory." We generated descriptions for patches in two different ways: (1) using only a given patch (i.e., 'exclusive' version), or (2) excluding the given patch (i.e., 'ablation' version). We considered the exclusive version as an initially more intuitive way for considering importance of each patch. Meanwhile, the ablation version was intended to mirror machine learning methods which determine importance of components in a model by selectively removing them (Hinton & Shallice, 1991; Meyes et al., 2019). As a control measure, we also generated a Qwen3-VL description for the entire unsegmented target image. All model descriptions were then transformed into text embeddings (i.e., "model embeddings") using MPNet.

Recall importance map analysis

We first considered whether competitor images may drive differences in verbal recall of the target image along specific patches. To do so, we created recall importance maps for each target image and each participant, then compared heatmaps between participants that studied the same competitor image versus other possible competitor images. We performed this analysis separately for the exclusive version and ablation version, and separately for each segmentation resolution (i.e., 2×2 to 5×5).

To create recall importance maps, we used model embeddings obtained from the Qwen3-VL patch descriptions. These model embeddings were generated for each patch in each target image segmentation (e.g., four model embeddings from four patches in 2×2 segmentation). We then computed the cosine similarity between model embeddings in each patch with recall embeddings of the target image, yielding an N×N recall importance maps for each participant and at each segmentation resolution (e.g., Fig. 3.1, right). Higher similarity values in a given patch indicate that patch was more important in verbal recall for the exclusive version, but less important in the ablation version. All similarity values were Fisher Z transformed. Finally, we considered whether importance maps varied based on the specific competitor image studied by computing Pearson's correlations between heatmaps from participants who either had (1) the same competitor image, or (2) other possible competitor images. For each participant, we computed a mean correlation for same-competitor pairs and other-competitor pairs. Finally, we ran paired-sample t-tests to assess whether heatmaps were more consistent for the same competitor versus other competitors (indicative of a competitor-specific influence on verbal recall).

For heatmap visualization, we performed two adjustments for interpretability: (1) baseline-correcting similarity values in the ablation version, and (2) controlling for target-driven similarity. We sought to baseline-correct values in the ablation version because lower correlations for a given patch indicated higher importance (opposite to exclusive version). As such, we defined a baseline constant as the cosine similarity between the target recall embedding and the model embedding for the entire unsegmented target image. Then, we subtracted similarity values in the ablation version from that constant. This adjustment allowed for patches with higher similarity to reflect patches that are more important. Next, for both exclusive and ablation versions, we controlled for target-driven similarity because target recall will be highly similar across participants regardless of the competitor studied. This overlap can make competitor-driven differences harder to visualize. As such, we first computed average importance maps for each competitor image, such that we had five heatmaps for each target (one from each competitor image). Next, for each competitor heatmap, we computed an across-competitor average heatmap that excluded itself (i.e., average heatmap across other four competitors). This across-competitor heatmap was computed to reflect target-driven similarity (i.e., patches that are important regardless of which competitor was studied). Finally, we subtracted the across-competitor heatmap from the held-out competitor's heatmap, effectively controlling for target-driven similarity. Together, these adjustments ensured (1) patch importance values reflected the same information in both versions, and (2) heatmaps reflected competitor-specific rather than target-driven patterns. We plotted these adjusted heatmaps for each category, version, competitor, and resolution.



Figure 3.1. Recall importance map analysis. Example 5×5 image segmentation (left). Cosine similarity computed between generated description text embeddings for each patch and recall text embedding for that image, yielding an importance heatmap (e.g., right).

Patch importance based on VGG-16 similarity

To determine whether importance of a given image patch was related to objective similarity of that patch across overlapping scenes, we compared similarity of image patches using VGG-16. For example, if a bright red car in the bottom right corner of the competitor image was driving the importance of an image patch in the bottom right corner, we might expect higher dissimilarity between the target and competitor image in that patch since the distinctive red car is absent in the target image. As such, we used feature representations obtained from three different layers in the VGG-16 model. These representations were obtained for each patch in each image segmentation (e.g., four representations from four patches in 2×2 segmentation). Next, we computed Pearson's correlations between representations for each target image with representations for each competitor image from the same category, separately for each segmentation resolution and separately for each VGG-16 layer. This yielded 'objective' VGG-16 similarity maps between the target and each of its possible competitor images. Correlation values were Fisher Z transformed.

Next, we considered whether these VGG-16 similarity maps corresponded with the recall importance maps obtained in the previous analysis. To do so, we first performed the same adjustments as before (baseline correcting similarity values in the ablation version and controlling for target-driven similarity in recall importance maps and VGG-16 similarity maps). Specifically, we subtracted all ablation version similarity values from the baseline constant (similarity between target recall and unsegmented generated description). Then, we accounted for target-driven influences in importance maps by computing an average across-competitor heatmap that excluded heatmaps from a held-out competitor, and subtracting the across-competitor heatmap from heatmaps for the held-out competitor. We performed this same adjustment to VGG-16 similarity maps by computing an average across-competitor objective similarity map that excluded heatmaps from a held-out competitor, and subtracting the across-competitor heatmap from the heatmap for the held-out competitor. Finally, for each version, segmentation resolution, and VGG-16 layer, we correlated VGG-16 similarity maps with recall importance maps and ran one-sample t-tests against zero. Significantly positive correlations would indicate that patches with higher objective similarity between the target and competitor were also more important in verbal recall of the target image. Meanwhile, significantly negative correlations would indicate that patches with higher objective similarity between the target and competitor were *less* important in verbal recall of the

target image. Or put another way, more dissimilar patches of the target and competitor were more important in recall of the target image.

Results

Verbal recall prioritizes competitor-specific image patches

We first generated recall importance maps for patches within the target image, based on the competitor it was paired with. Heatmaps for both the exclusive and ablation versions, and for each segmentation resolution are shown (Figs. 3.2-7). Warmer colors indicate higher importance for a given patch and cooler colors indicate lower importance. Next, we compared importance maps between participants that had either (1) the same competitor, or (2) other possible competitors, using paired-sample t-tests. Higher correlations between heatmaps for the same competitor would indicate that each competitor influences verbal recall of the target along specific and consistent patches. In the exclusive version of the analysis, we found that same competitor heatmaps were significantly more similar than other competitor heatmaps at only the 5×5 resolution (2×2: $t_{(89)} = 0.90$, $p = 0.37$; 3×3: $t_{(89)} = 0.46$, $p = 0.64$; 4×4: $t_{(89)} = 0.93$, $p = 0.35$; 5×5: $t_{(89)} = 3.29$, $p = 0.001$; Fig. 3.8a). Meanwhile, in the ablation version of the analysis, we found that same competitor heatmaps were significantly more similar than other competitor heatmaps at the 3×3 and 5×5 resolution (2×2: $t_{(89)} = 1.82$, $p = 0.07$; 3×3: $t_{(89)} = 2.80$, $p = 0.006$; 4×4: $t_{(89)} = 1.93$, $p = 0.06$; 5×5: $t_{(89)} = 2.49$, $p = 0.01$; Fig. 3.8b). This finding demonstrates that verbal recall prioritizes different image patches based on the specific paired competitor image. Moreover, the stronger effects in the ablation version suggest ablation is a more sensitive approach for quantifying patch importance, consistent with its use in machine learning work.

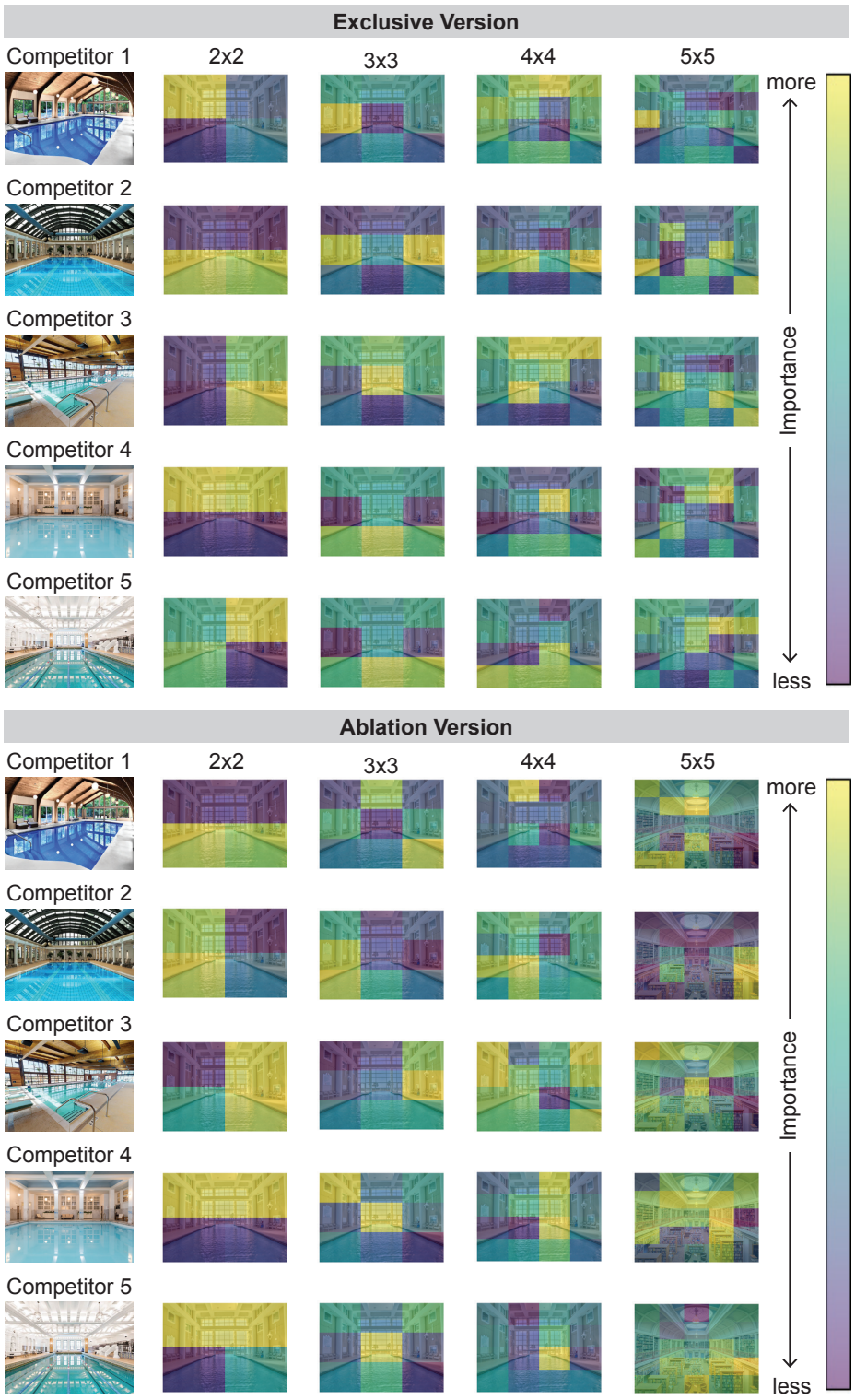

Figure 3.2. Pool recall importance maps. Recall importance maps for category pool, separately for each competitor and each segmentation resolution. Warmer colors indicate higher importance, and cooler colors indicate lesser importance. Exclusive version heatmaps above and ablation version heatmaps below.

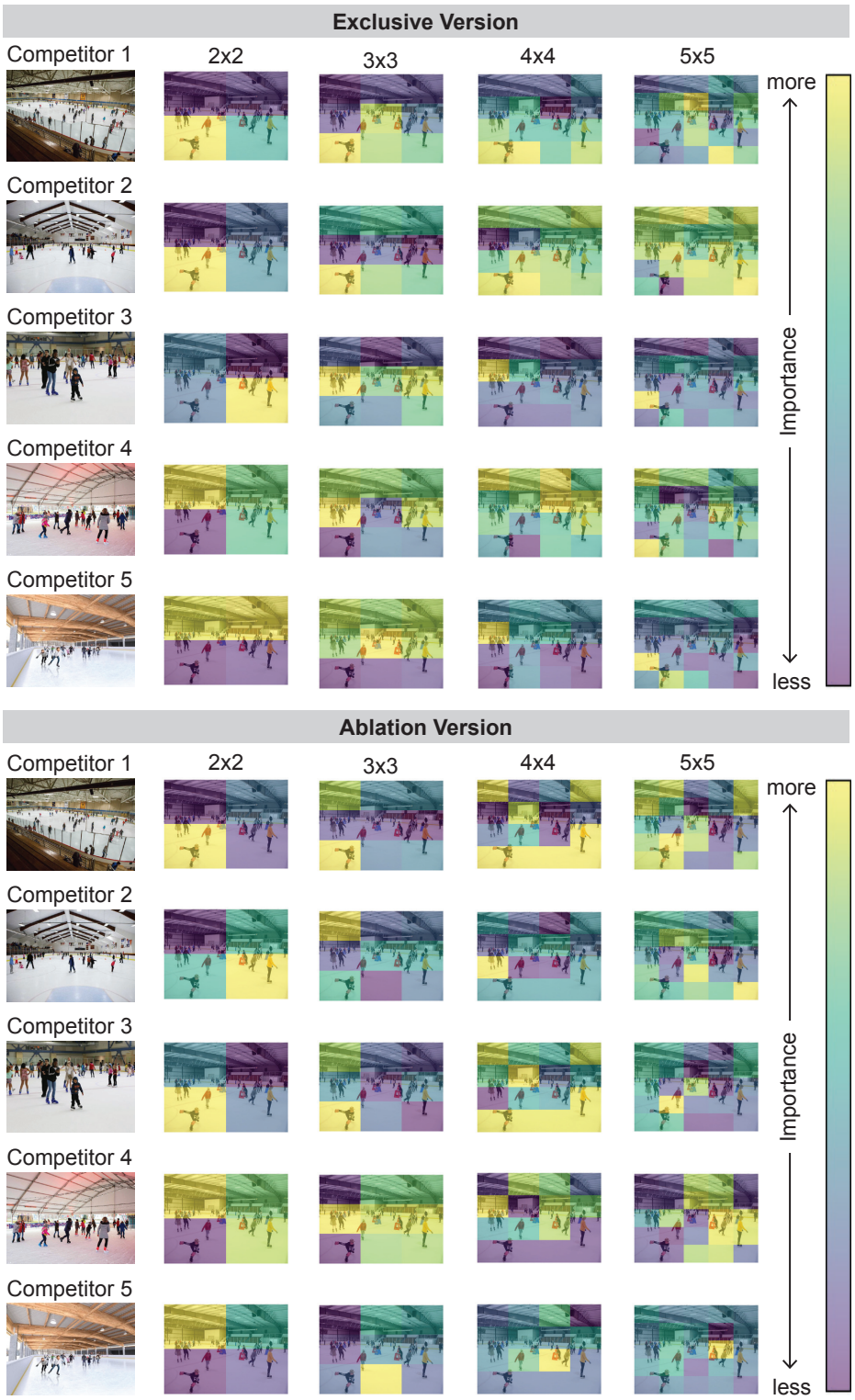

Figure 3.3. Skating recall importance maps. Recall importance maps for category skating, separately for each competitor and each segmentation resolution. Warmer colors indicate higher importance, and cooler colors indicate lesser importance. Exclusive version heatmaps above and ablation version heatmaps below.

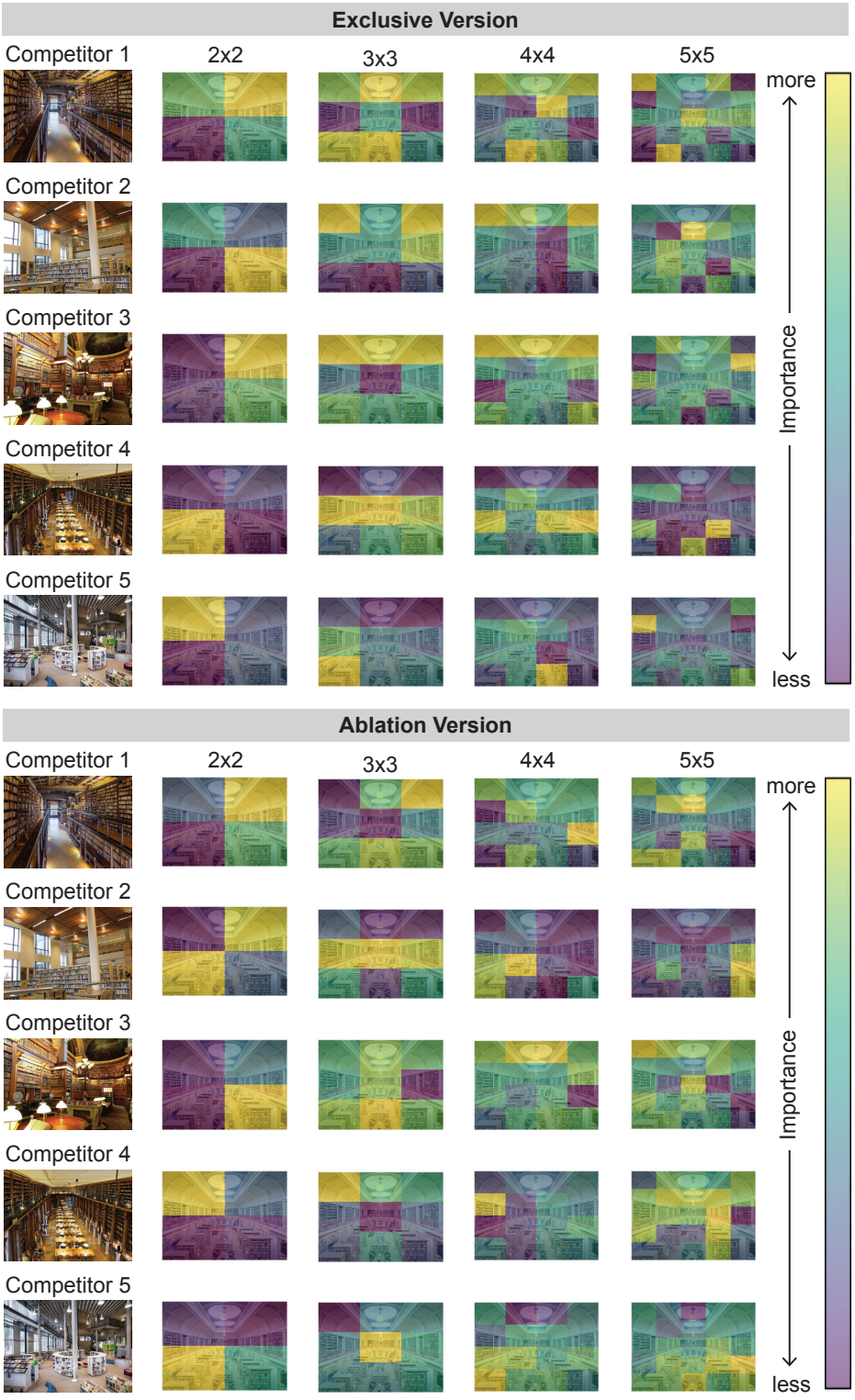

Figure 3.4. Library recall importance maps. Recall importance maps for category library, separately for each competitor and each segmentation resolution. Warmer colors indicate higher importance, and cooler colors indicate lesser importance. Exclusive version heatmaps above and ablation version heatmaps below.

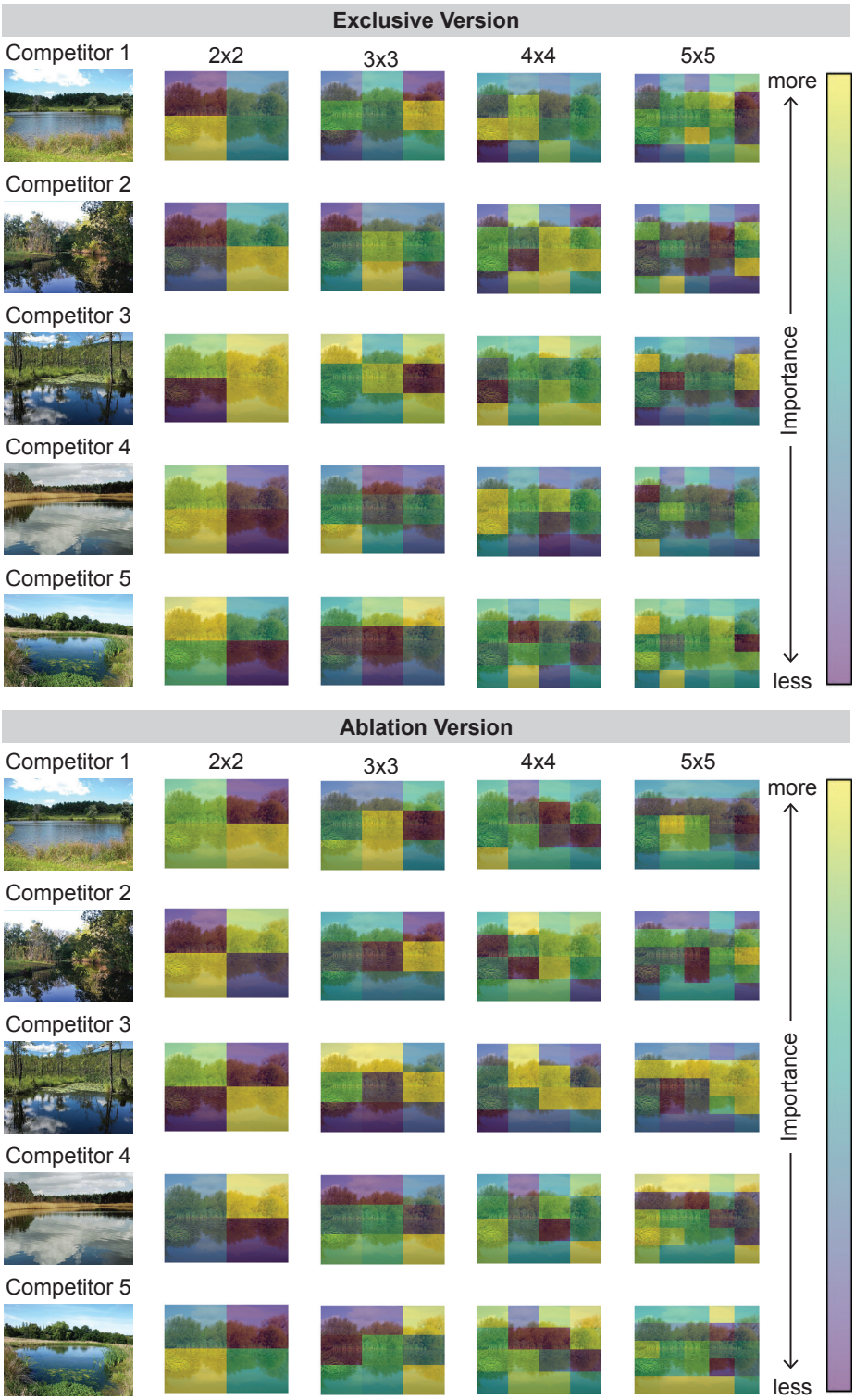

Figure 3.5. Pond recall importance maps. Recall importance maps for category pond, separately for each competitor and each segmentation resolution. Warmer colors indicate higher importance, and cooler colors indicate lesser importance. Exclusive version heatmaps above and ablation version heatmaps below.

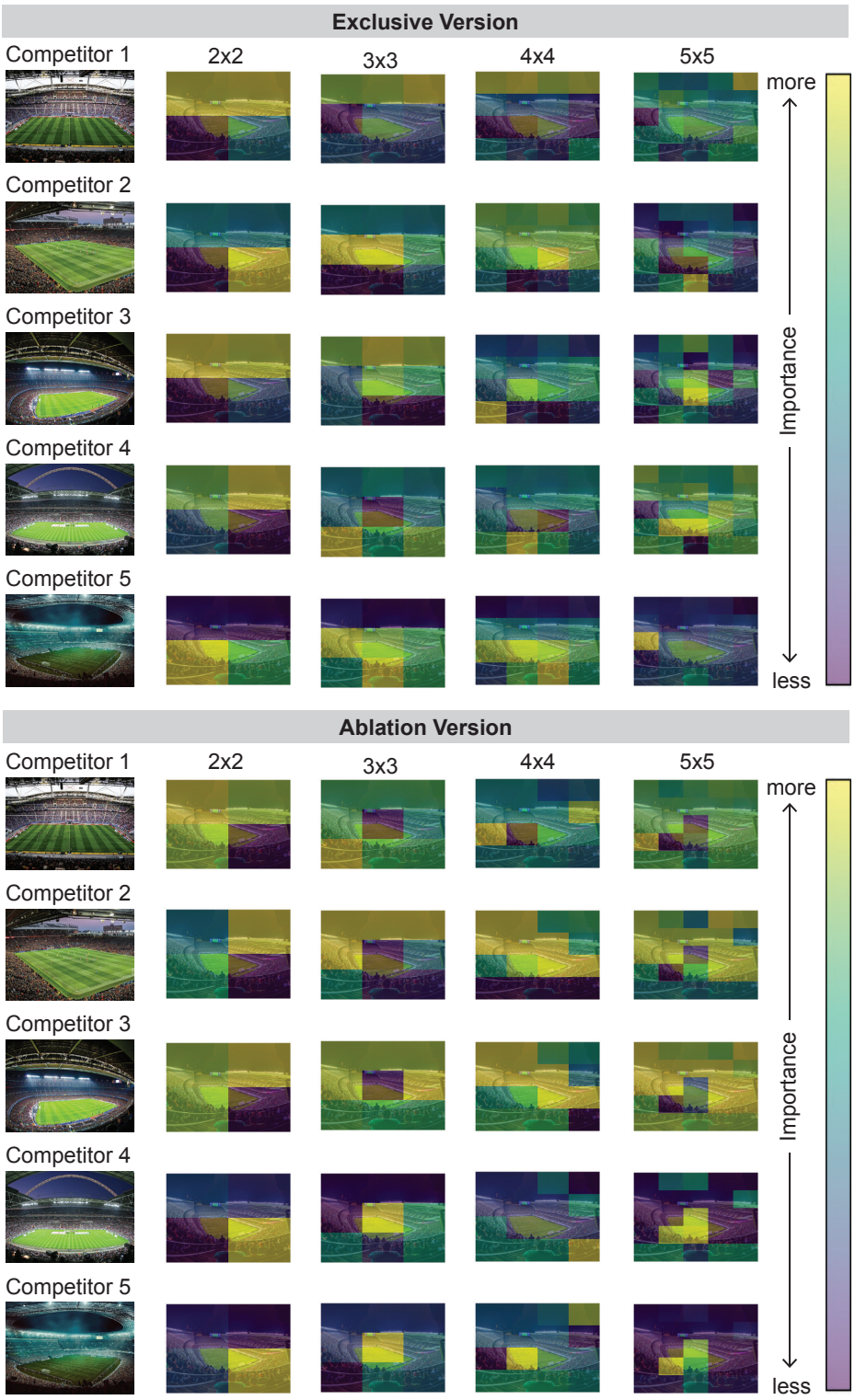

Figure 3.6. Stadium recall importance maps. Recall importance maps for category stadium, separately for each competitor and each segmentation resolution. Warmer colors indicate higher importance, and cooler colors indicate lesser importance. Exclusive version heatmaps above and ablation version heatmaps below.

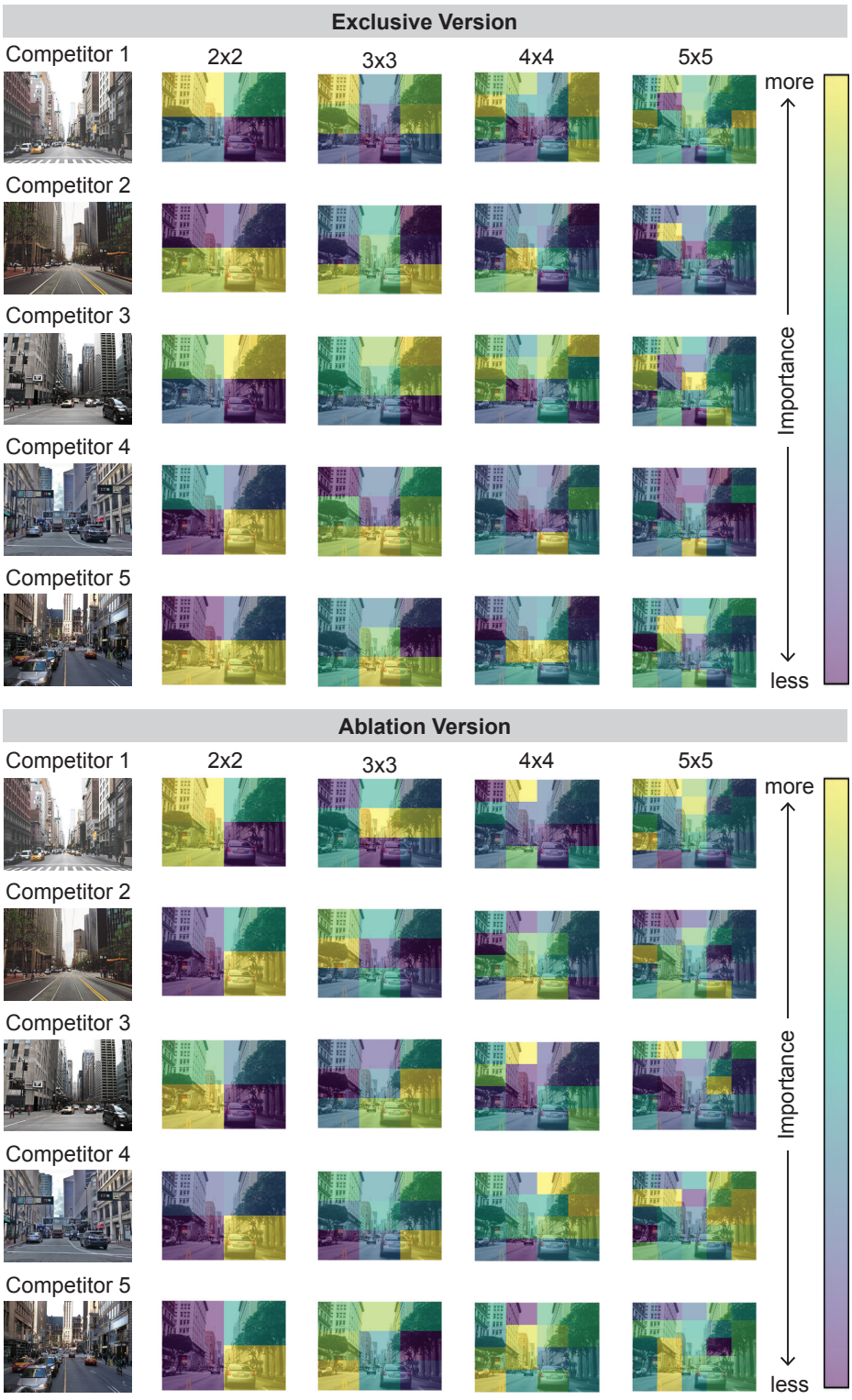

Figure 3.7. Downtown recall importance maps. Recall importance maps for category downtown, separately for each competitor and each segmentation resolution. Warmer colors indicate higher importance, and cooler colors indicate lesser importance. Exclusive version heatmaps above and ablation version heatmaps below.

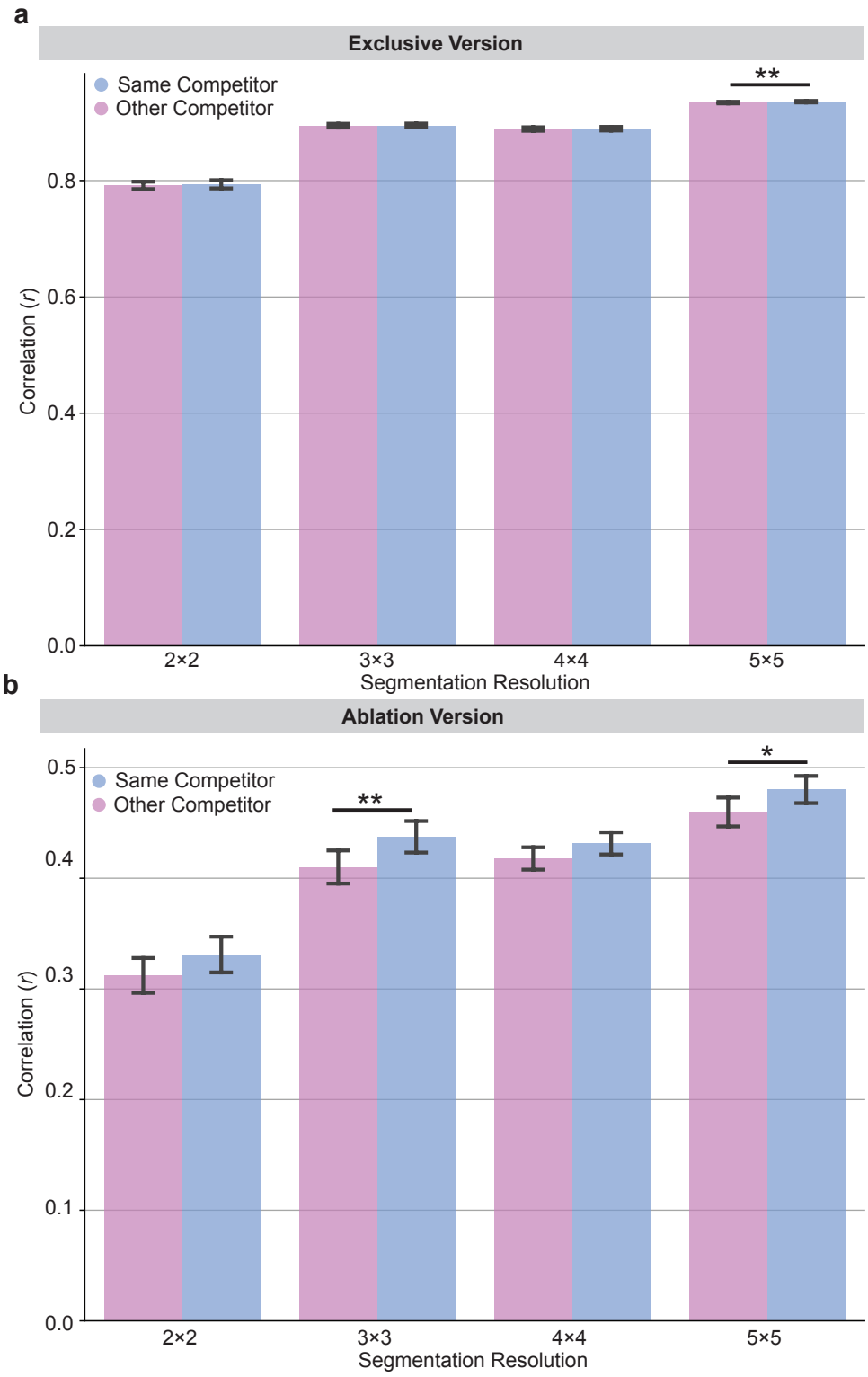

Figure 3.8. Competitor-specific recall importance maps. a) Exclusive version. Pearson's correlations between target image recall importance maps for either (1) the same competitor, or (2) other competitors. **b)** Ablation version. Pearson's correlations between target image recall importance maps for either (1) the same competitor, or (2) other competitors.

Patch importance driven by VGG-16 similarity

The first analysis indicated patch importance in target verbal recall depends on the specific competitor image studied. Next, we sought to determine whether patch importance was driven by objective similarity of that patch between the target and competitor. We predicted that objectively dissimilar patches would be more important for memory discrimination (and therefore prioritized in verbal recall). To do so, we compared VGG-16 similarity heatmaps from three VGG-16 layers to recall importance maps. We then compared the distribution of similarity values to zero using a one-sample t-test. If patch importance was driven by dissimilarity between the target and competitor, we would expect a significantly negative distribution.

In the exclusive version, we found that Layer 1 similarity was not significantly related to importance maps at any segmentation resolution (2×2: $t_{(89)} = 0.008$, $p = 0.99$; 3×3: $t_{(89)} = 0.44$, $p = 0.66$; 4×4: $t_{(89)} = 1.11$, $p = 0.27$; 5×5: $t_{(89)} = 0.70$, $p = 0.49$; Fig. 3.9a). Meanwhile, we found that Layer 3 similarity was negatively related to importance maps at 4×4 resolution (2×2: $t_{(89)} = 1.29$, $p = 0.20$; 3×3: $t_{(89)} = 1.73$, $p = 0.09$; 4×4: $t_{(89)} = 2.06$, $p = 0.04$; 5×5: $t_{(89)} = 0.70$, $p = 0.49$; Fig. 3.9a) and Layer 5 similarity was negatively related to importance maps at the 3×3 and 4×4 resolution (2×2: $t_{(89)} = 0.12$, $p = 0.90$; 3×3: $t_{(89)} = 3.29$, $p = 0.001$; 4×4: $t_{(89)} = 3.06$, $p = 0.003$; 5×5: $t_{(89)} = 1.14$, $p = 0.26$; Fig. 3.9a). In the ablation version, we found that Layer 1 similarity was negatively related to importance maps at 3×3, 4×4, and 5×5 resolution (2×2: $t_{(89)} = 0.44$, $p = 0.66$; 3×3: $t_{(89)} = 2.53$, $p = 0.01$; 4×4: $t_{(89)} = 2.02$, $p = 0.046$; 5×5: $t_{(89)} = 3.35$, $p = 0.001$; Fig. 3.9b). Meanwhile, Layer 3 similarity was negatively related to importance maps at 3×3 resolution (2×2: $t_{(89)} = 1.51$, $p = 0.14$; 3×3: $t_{(89)} = 2.59$, $p = 0.01$; 4×4: $t_{(89)} = 0.38$, $p = 0.70$; 5×5: $t_{(89)} = 1.42$, $p = 0.16$; Fig. 3.9b) and Layer 5 similarity was negatively related to importance maps at the 5×5 resolution (2×2: $t_{(89)} = 0.03$, $p = 0.97$; 3×3: $t_{(89)} = 0.02$, $p = 0.98$; 4×4: $t_{(89)} = 1.43$, $p = 0.16$; 5×5: $t_{(89)} = 2.49$, $p = 0.01$; Fig. 3.9b). These results, while varied, reveal a common pattern: VGG-16 similarity between target and competitor was negatively associated with recall importance at a given patch. In other words, dissimilarity between specific patches of the target and competitor seems to drive verbal recall of the target image. Moreover, these effects appear specific to higher segmentation resolutions (3×3, 4×4, 5×5), suggesting the organization of features at a lower resolution like 2×2 is too coarse to capture subtle competitor-specific biases in verbal recall.

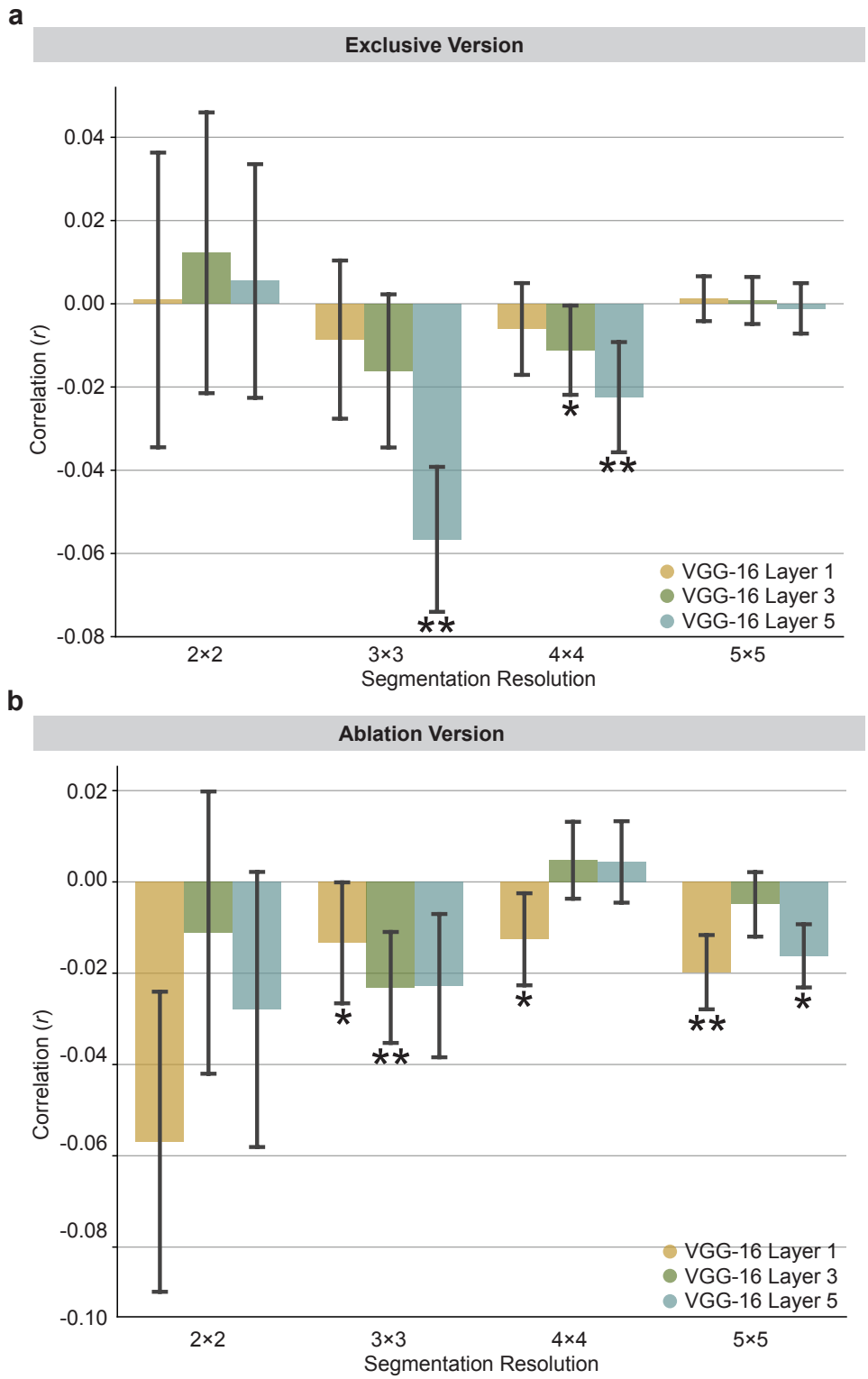

Figure 3.9. VGG-16 similarity predicts recall importance. a) Exclusive version. Distributions of Pearson's correlations between target image recall importance maps and VGG-16 similarity maps. **b)** Ablation version. Distributions of Pearson's correlations between target image recall importance maps and VGG-16 similarity maps.

General Discussion

In this chapter, we develop an 'image patch' analysis to consider whether specific patches of a scene image drive distinctiveness in verbal recall. Using this method, we demonstrate that each competitor image induced a unique and consistent profile of patch importance in verbal recall. Moreover, we found that dissimilarity between the target and competitor image at different patches drives competitor-specific differences in patch importance. Taken together, we provide evidence for how overlapping scene images affect verbal recall – by prioritizing dissimilar patches of overlapping images.

Most studies of visual memory require separate retrieval tasks in order to probe memory for multiple features, leading to the use of simple and highly controlled stimuli (Bae & Luck, 2017; Chanales et al., 2021; Chapman et al., 2023; Chunharas et al., 2022; Golomb, 2015; Magnussen & Greenlee, 1999; Morrow & Clewett, 2025; Rademaker et al., 2015). Recently, however, verbal recall descriptions have allowed memory for naturalistic image and movie stimuli to be quantified across hundreds of dimensions from a single retrieval task (Chen et al., 2017; Wang et al., 2023). However, verbal recall quantified via text embeddings is difficult to interpret and, as we observe in Chapter II, can lead to ambiguous results. In this chapter, we create recall importance maps to translate verbal recall into a more interpretable 2D feature space. A key benefit of this analysis is that we can maintain richness of information in memory content *and* avoid explicitly probing memory for each image patch.

In our analysis, we translate verbal recall into an explicit feature space (i.e., image patches) without explicitly probing memory for patches, allowing us to characterize which patches of a scene contribute to distinctiveness in verbal recall. This approach mirrors visual attention studies using eye-tracking, which show that fixations during encoding of visual stimuli determine what is later remembered (Loftus, 1972; Pertzov et al., 2009). By mapping verbal recall to different patches of an image, our importance heatmaps are analogous to scanpath analyses in eye-tracking studies (Anderson et al., 2015; klein Selle et al., 2021). However, the 2D space of images is not the only space we could use to translate verbal recall into explicit features; for example, some large language models output explicit feature tags such as valence, arousal, concreteness, etc. (Martínez et al., 2024). These tags could then be used to quantify explicit features along which verbal recall for overlapping images varies. In this manner, our approach could be extended to consider many other features which contribute to distinctiveness in verbal recall for overlapping images.

In summary, we demonstrate that verbal recall for overlapping images is anchored to specific patches, and that selection of these patches is driven by objective dissimilarity between the overlapping images in those regions. By developing our image patch analysis, we show that unconstrained and high-dimensional data such as verbal recall can be translated into explicit and interpretable feature spaces. Future work can expand upon these methods to decompose verbal recall into other dissociable components. A promising avenue of work is to consider which aspects of verbal recall for overlapping images are common across individuals versus idiosyncratic. As such, we present the current chapter as a proof of concept for a new method, which can provide insight into how overlap adaptively shapes verbal recall.

References Cited

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., … Zheng, X. (2016). *TensorFlow: A system for large-scale machine learning* (arXiv:1605.08695). arXiv. https://doi.org/10.48550/arXiv.1605.08695

Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2015). A comparison of scanpath comparison methods. *Behavior Research Methods*, *47*(4), 1377–1392. https://doi.org/10.3758/s13428-014-0550-3

Bae, G.-Y., & Luck, S. J. (2017). Interactions between visual working memory representations. *Attention, Perception & Psychophysics*, *79*(8), 2376–2395. https://doi.org/10.3758/s13414-017-1404-8

Bai, S., Cai, Y., Chen, R., Chen, K., Chen, X., Cheng, Z., Deng, L., Ding, W., Gao, C., Ge, C., Ge, W., Guo, Z., Huang, Q., Huang, J., Huang, F., Hui, B., Jiang, S., Li, Z., Li, M., … Zhu, K. (2025). *Qwen3-VL Technical Report* (arXiv:2511.21631). arXiv. https://doi.org/10.48550/arXiv.2511.21631

Chanales, A. J. H., Tremblay-McGaw, A. G., Drascher, M. L., & Kuhl, B. A. (2021). Adaptive Repulsion of Long-Term Memory Representations Is Triggered by Event Similarity. *Psychological Science*, *32*(5), 705–720. https://doi.org/10.1177/0956797620972490

Chapman, A. F., Chunharas, C., & Störmer, V. S. (2023). Feature-based attention warps the perception of visual features. *Scientific Reports*, *13*(1), 6487. https://doi.org/10.1038/s41598-023-33488-2

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Chollet, F. (2015). *Keras* [Computer software]. https://github.com/fchollet/keras

Chunharas, C., Rademaker, R. L., Brady, T. F., & Serences, J. T. (2022). An adaptive perspective on visual working memory distortions. *Journal of Experimental Psychology. General*, *151*(10), 2300–2323. https://doi.org/10.1037/xge0001191

Drascher, M. L., & Kuhl, B. A. (2022). Long-term memory interference is resolved via repulsion and precision along diagnostic memory dimensions. *Psychonomic Bulletin & Review*, *29*(5), 1898–1912. https://doi.org/10.3758/s13423-022-02082-4

Golomb, J. D. (2015). Divided spatial attention and feature-mixing errors. *Attention, Perception, & Psychophysics*, *77*(8), 2562–2569. https://doi.org/10.3758/s13414-015-0951-0

Hinton, G. E., & Shallice, T. (1991). Lesioning an attractor network: Investigations of acquired dyslexia. *Psychological Review*, *98*(1), 74–95. https://doi.org/10.1037/0033-295x.98.1.74

klein Selle, N., Gamer, M., & Pertzov, Y. (2021). Gaze-pattern similarity at encoding may interfere with future memory. *Scientific Reports*, *11*(1), 7697. https://doi.org/10.1038/s41598-021-87258-z

Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, *3*(4), 525–551. https://doi.org/10.1016/0010-0285(72)90021-7

Magnussen, S., & Greenlee, M. W. (1999). The psychophysics of perceptual memory. *Psychological Research*, *62*(2), 81–92. https://doi.org/10.1007/s004260050043

Martínez, G., Molero, J. D., González, S., Conde, J., Brysbaert, M., & Reviriego, P. (2024). Using large language models to estimate features of multi-word expressions: Concreteness, valence, arousal. *Behavior Research Methods*, *57*(1), 5. https://doi.org/10.3758/s13428-024-02515-z

Meyes, R., Lu, M., Puiseau, C. W. de, & Meisen, T. (2019). *Ablation Studies in Artificial Neural Networks* (arXiv:1901.08644). arXiv. https://doi.org/10.48550/arXiv.1901.08644

Morrow, E., & Clewett, D. (2025). Distortion of overlapping memories relates to arousal and anxiety. *Cognitive, Affective & Behavioral Neuroscience*, *25*(1), 154–172. https://doi.org/10.3758/s13415-024-01229-1

Pertzov, Y., Avidan, G., & Zohary, E. (2009). Accumulation of visual information across multiple fixations. *Journal of Vision*, *9*(10), 2.1-12. https://doi.org/10.1167/9.10.2

Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(6), 1650–1665. https://doi.org/10.1037/xhp0000110

Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks* (arXiv:1908.10084). arXiv. https://doi.org/10.48550/arXiv.1908.10084

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv:1409.1556). arXiv. https://doi.org/10.48550/arXiv.1409.1556

Song, K., Tan, X., Qin, T., Lu, J., & Liu, T.-Y. (2019). *MASS: Masked Sequence to Sequence Pre-training for Language Generation*. https://doi.org/10.48550/ARXIV.1905.02450

Tandoc, M. C., Dong, C. V., & Schapiro, A. C. (2024). Object Feature Memory Is Distorted by Category Structure. *Open Mind: Discoveries in Cognitive Science*, *8*, 1348–1368. https://doi.org/10.1162/opmi_a_00170

Wang, Y., Lee, H., & Kuhl, B. A. (2023). *Mapping multidimensional content representations to neural and behavioral expressions of episodic memory* [Preprint]. Neuroscience. https://doi.org/10.1101/2023.01.09.523096

CHAPTER IV
NEURAL ENCODING PREDICTS SHARED VERSUS IDIOSYNCRATIC VERBAL RECALL

The design and analyses described in this chapter were developed by a number of lab members, including Anisha S. Babu, Zhifang Ye, Julian Gamez, and Brice A. Kuhl. Zhifang Ye and Brice A. Kuhl contributed substantially by aiding in the development of fMRI analysis methods. Julian Gamez aided in data collection and preliminary VGG-16 analyses. I was the primary contributor for data collection, data analysis, and writing.

Introduction

In cognitive neuroscience studies of human memory, one of the most influential experimental approaches is to consider whether neural activity during memory encoding predicts subsequent remembering (Gordon et al., 2014; H. Kim, 2011; Staresina et al., 2011; Wagner et al., 1999; Watanabe et al., 2011). These approaches generally define subsequent memory in terms of objective, categorical outcomes (e.g., remembered vs. forgotten). However, memory outcomes are often much more nuanced. For example, even when two individuals are successfully remembering the exact same event, the contents of their memories are likely to differ slightly (Bartlett, 1932; Chen et al., 2017; Heusser et al., 2021). In other words, even successful remembering can contain idiosyncratic content. Thus, an important extension of classic subsequent memory effects is to identify neural measures that predict idiosyncratic content in subsequent expressions of memory.

To date, human neuroimaging studies investigating predictors of shared versus idiosyncratic memory content have primarily focused on (1) intersubject correlations of activity patterns, or (2) alignment of representational structure across individuals. For example, several studies have found that when different individuals are encoding the same event, neural activity patterns are positively correlated across individuals (intersubject correlations; Chen et al., 2017; Hasson et al., 2004; Jääskeläinen et al., 2008). Moreover, higher intersubject correlations are associated with greater similarity in the *content* of subsequent memories (Nguyen et al., 2019; Sava-Segal et al., 2023). Or put another way, greater neural *dissimilarity* between subjects is associated with more idiosyncratic remembered content. Conceptually similar results have been found when correlating the *representational structure* of events across individuals (as opposed to directly correlating activity patterns; Sheng et al., 2023). Collectively, these studies provide

important and compelling evidence linking neural measures of memory encoding to subject-specific differences in remembered content.

While measures of intersubject similarity have notable strengths, there are also some limitations to this approach. A primary limitation is that, on its own, neural *dissimilarity* across subjects can be explained in multiple ways (Finn et al., 2020). The most appealing account, of course, is that dissimilarity between subject 1 and subject 2 reflects idiosyncratic (but otherwise successful) encoding (Hasson et al., 2009). An alternative, however, is that dissimilarity reflects noise—either measurement noise or noise due to psychological factors (Finn et al., 2020). Of particular concern, noise can be driven by disengagement from the task (Ki et al., 2016; H. Song et al., 2021). For example, if subject 1 is disengaged during encoding whereas subject 2 is highly engaged, neural encoding is likely to be relatively dissimilar. Moreover, disengagement can also explain why subject 1 and subject 2 would have correspondingly low similarity in their recalled content: namely, subject 1 failed to encode what subject 2 successfully encoded. Thus, neural dissimilarity does not necessarily indicate that each subject successfully encoded something different; rather, it may reflect that one subject encoded more successfully than the other subject. Indeed, higher intersubject correlations (potentially reflecting synchronization with engagement) are associated with better subsequent memory (Hasson et al., 2008; Sheng et al., 2023). While this account is not likely to fully explain previously reported evidence of idiosyncratic encoding measured by intersubject correlations, the key point is that intersubject correlations are open to multiple interpretations.

In this chapter, we sought to more directly link neural measures of memory encoding to subsequent expressions of idiosyncratic memory recall. The key innovation in our approach is that, rather than correlating neural measures of encoding across individuals, we focused on the *within-participant* relationship between the representational structure of encoding and recall content. Specifically, using fMRI, we measured each participant's representational structure for a set of natural scene images during memory encoding. Then, we measured the representational structure of those same stimuli when they were subsequently recalled from memory (using large language models applied to verbal recall). This allowed us to compare each participant's neural encoding to (a) *their own verbal recall* and (b) recall content shared across *other participants* (i.e., 'normative' recall content). To preview, we show that neural measures of memory encoding indeed reflect participant-specific (idiosyncratic) details of subsequent verbal recall beyond what was explained

by normative recall content. These neural measures of idiosyncratic encoding were most prominent in visual cortical areas. In contrast, areas of parietal cortex preferentially reflected normative recall content. Finally, the hippocampus exhibited a unique pattern of results: pattern similarity in the hippocampus was *negatively* related to participant-specific (idiosyncratic) measures of subsequent recall similarity, but positively related to normative recall content. In other words, hippocampal activity patterns during encoding were most *dissimilar* when participant -specific recall content was most similar. Collectively, these results demonstrate that neural measures of memory encoding are sensitive to both normative and idiosyncratic details of recalled content, with this sensitivity varying sharply across brain regions.

Methods

Participants

A target sample size of 40 participants was determined in advance. To reach this sample, forty-four healthy participants were recruited from the University of Oregon and broader community. All participants self-reported to be right-handed, native English speakers, have normal or corrected-to-normal vision and hearing, and no history of neurologic disorders. Informed consent was obtained from all participants according to a protocol approved by the University of Oregon Institutional Review Board, and all participants were paid for their participation ($25/hr fMRI and $10/hr behavioral). The first two participants were excluded due to technical issues with saving behavioral data. Two participants were excluded because they failed to reach a pre-determined behavioral threshold (see 'Experimental procedure'). Two participants were excluded for excessive head movement (see 'fMRI preprocessing'). We therefore report results for 38 participants (24 female; mean age 20.92 ± 3.57 SD, range 18-34).

Stimuli

Thirty-six scene images and 18 face images were used in the experiment. The 36 scene images were drawn equally from six different categories: indoor pool, ice skating rink, library, pond, stadium, and downtown street (Fig. 4.1a). Critically, these categories were separated into indoor (former three) and outdoor (latter three) groups. Participants were alternately assigned to either the indoor or outdoor group, and data was collapsed across groups for analysis. Each of the 18 scene images assigned to a participant were then randomly paired with a face image, creating a

set of 18 face-scene paired associates. The faces served as cues during cued recall tasks (described below).

Additionally, we defined visual features in each stimulus using convolutional neural network VGG-16 (Simonyan & Zisserman, 2015). First, each scene image was converted into array format and inputted into VGG-16 using TensorFlow and Keras libraries (Abadi et al., 2016; Chollet, 2015). Next, we extracted features from an early, middle, and late layer of processing in VGG-16 (block 1 pool, block 3 pool, block 5 pool). This yielded 802814-, 200702-, and 25088-dimensional vectors, respectively, representing features that vary from more visual (early layers) to more semantic (late layers).

Experimental procedure

Data collection occurred across two days (separated on average by 5.63 days, range 1-29 days). After providing consent, participants entered the MRI scanner. On the first day, participants provided informed consent, and we obtained anatomical scans (~15min). Additional functional scans were obtained for unrelated tasks (to be reported elsewhere). On the second day, participants provided informed consent and reviewed instructions. Inside the scanner, participants completed six functional runs of the experimental paradigm (Fig. 4.1b). Each run included two phases: encoding and retrieval. The start of each phase was preceded by a screen introducing the phase as "Study" or "Test" (i.e., encoding or retrieval; 7500ms). Across all phases, stimuli were displayed on a white background, projected from the back of the scanner. After exiting the scanner, participants completed two additional behavioral memory tasks. The experiment was implemented in PsychoPy (Peirce et al., 2019) and jsPsych (Leeuw et al., 2023) hosted on Pavlovia (https://pavlovia.org), and lasted ~2h30min, with ~1h45min inside the scanner.

During the encoding phases, participants learned 18 face-scene associations, one association at a time. Phases were separated into two blocks, with each association presented once within each block. Each encoding trial began with the presentation of a face image (1000ms), followed by a black fixation cross (500ms), the associated scene image (3000ms), then another black fixation cross (500ms). Between each trial, participants completed a simple engagement task where they reported the direction of an arrow (left/right) via button press. A black arrow as presented (1000ms) followed by another fixation cross (1500ms) until the start of the next trial.

The order in which the 18 face-scene associations were encoded was randomized within each run and block, and for each participant.

During the retrieval phases, participants were cued to recall the scene associated with each of the 18 faces, one association at a time. Phases were separated into two blocks, with each association cued once within each block. Each retrieval trial began with a face image (1000ms), followed by a black fixation cross (500ms), then a white box (1000ms) as a cue to retrieve the associated scene image. When the white box turned red (1500ms), participants were asked to report via button press the vividness of their memory on a three-point scale ("no memory", "weak memory", "vivid memory"). A black fixation cross (1500ms) was then presented until the start of the next trial. The order in which the 18 face-scene associations were retrieved were randomized within each run and block, and for each participant.

After exiting the scanner, participants completed a cued recall test. For each face-scene associate, participants were shown a cue (face) at the top of the screen and were asked to type a description of the associated scene image (Fig. 4.1c). Specifically, participants were asked to complete a description beginning with: "An image of …" Further, they were asked to avoid describing the face image, using unnecessary words (e.g., "an image of"), or referencing other scene images. Example good and bad descriptions were included in the instructions. After participants typed their response on the computer screen, they pressed 'continue' to advance to the next trial. No time limit was given, though participants were required to use at least 10 words in their descriptions in order to proceed. Further, if participants did not remember the associated scene image, they could check a box indicating that they did not remember and were forced to wait 30s before proceeding to the next trial.

Following cued verbal recall, participants completed a multiple-choice associative memory test for each of the 18 face-scene associates. Here, participants were presented with each face image at the top of the screen and six scene images below. Critically, these images consisted of two images from each of the three studied categories (e.g., two pools, two ice skating rinks, two libraries). Same-category and across-category lure images were randomly selected. Participants were asked to click on the correct scene image using the mouse, then click on their confidence (sure/guess). High accuracy in this test was rewarded with a cash bonus (chance accuracy = 16.67%). Participants missing >1 verbal recall response and/or incorrectly responding to >1 multiple choice question were excluded from analysis (n=2 excluded). Of the remaining N=38

participants, N=6 participants were missing one verbal recall response and N=6 participants missed one multiple choice question (overlap N=1).

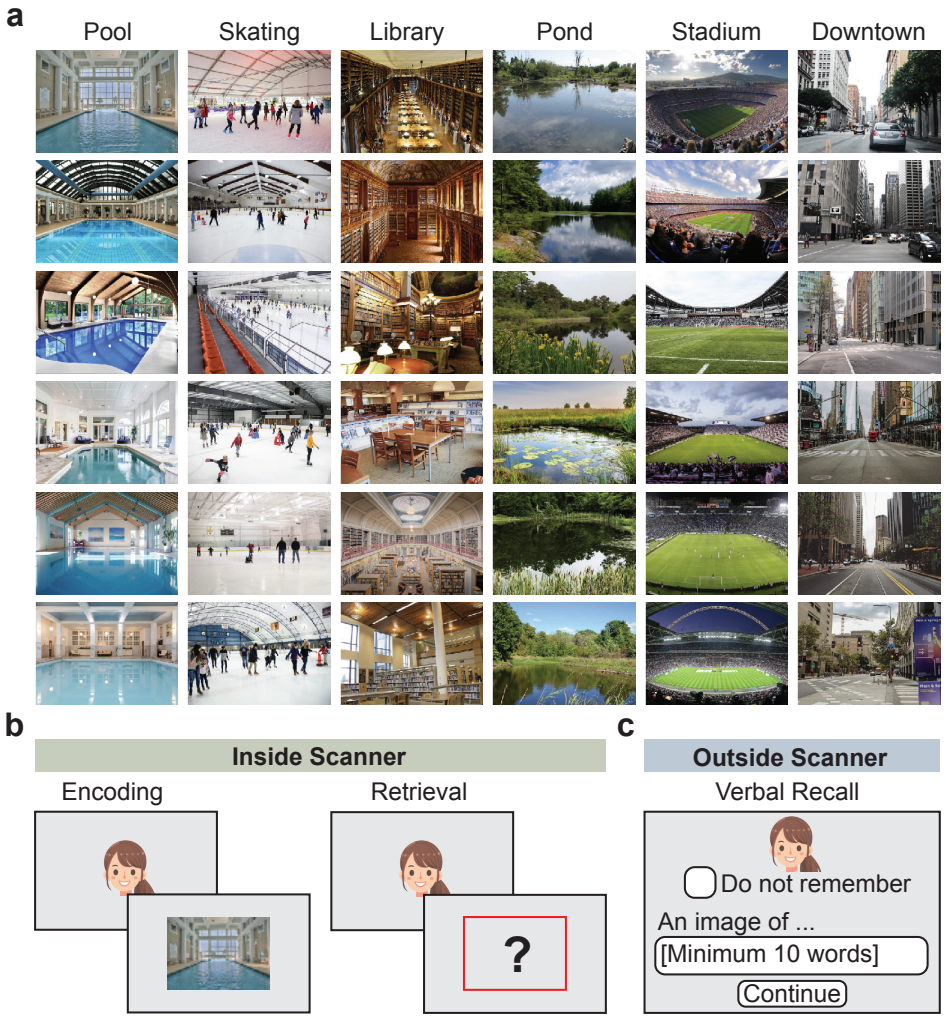

Figure 4.1. Stimuli & task design. a) Stimuli. Participants were assigned all images from three categories (i.e., first three columns or second three columns). **b)** Learning task (inside fMRI scanner). Learning consisted of encoding and retrieval trials. **c)** Recall task (outside fMRI scanner). After learning, participants were cued with each face image at the top of the screen and asked to type a description of the associated scene image.

Analysis of behavioral data

Vividness ratings

In each of the six retrieval phases, participants rated the vividness of their memory for each scene image on a 3-point scale ("no memory", "weak memory", "vivid memory"). Vividness ratings were recoded to a numerical scale (1-3). To compare ratings across runs, we applied a

mixed effects model using the lme4 package in R to predict mean vividness rating as a function of learning run with participant as a random effect: vividness_rating ~ learning_run + (1|participant).

Verbal recall

After learning runs in the scanner, participants described each scene image from memory. To determine if there were any effects of image category, we compare the mean number of words in verbal recall for each category using a one-way ANOVA with factor category, separately for 'indoor' and 'outdoor' categories. Additionally, we tested for differences between 'indoor' and 'outdoor' category participants using a two-sample t-test.

Next, we defined features in verbal recall obtained from the cued verbal recall task by transforming descriptions into text embeddings using large language model MPNet (Fig. 4.3a). Missing descriptions (maximum one/subject) were coded as null for analysis. First, we removed punctuation, extra whitespace, and changed all characters to lowercase using the Gensim preprocessing library (Rehurek, 2011). Next, we transformed each description into a text embedding using MPNet (i.e., 'paraphrase-mpnet-base-v2' from the Sentence Transformers library; Reimers & Gurevych, 2019; K. Song et al., 2019). This yielded 768-dimensional vectors representing semantic information in verbal descriptions for each image and each participant. Text embeddings from each participant were used to represent multidimensional features in memory content. We plot the relative position of text embeddings in two-dimensional space using multidimensional scaling (MDS) based on Euclidean distance using the function dist() from the stats package in R (Fig. 4.3b). To visualize the category structure, embeddings are colored by the category of the image they are meant to describe.

For an analysis comparing participant-specific and normative features in recall, we also constructed averaged text embedding across participants for each image. Critically, average text embeddings were calculated separately for each participant such that a participant's own recall would not be included in the average. These averaged embeddings are referred to as normative (across-subject) recall.

Representational similarity analysis

We generated similarity matrices using cosine similarity for VGG-16, and verbal recall. We chose to use cosine similarity which prioritizes directional differences between vectors over

Pearson correlation, though we do not suspect significant differences had Pearson correlation been used instead. A single VGG-16 similarity matrix was calculated for each of the three selected layers. Meanwhile, participant-specific verbal recall similarity was calculated separately for each participant (38 recall similarity matrices; e.g., Fig. 4.3c). Normative across-subject recall similarity was also calculated separately for each subject (38 recall similarity matrices). Next, a Fisher z-transformation was applied to each value in each similarity matrix. Then, we extracted only within-category comparisons from each matrix. To provide a fair comparison in models including multiple predictors, we z-scored values within the VGG-16 matrix, across all participant-specific recall matrices, and across all normative recall matrices.

Category sensitivity in recall

To better illustrate within-category recall similarity in each of the six categories, we plotted the distribution of similarity values in each category, with a red line representing average across-category similarity for images within that category (e.g., pool2-library3 similarity included in both pool and library across-category averages; Fig. 4.3d). We compared within- and across-category similarity in each category by running a one-sample t-test for each category using mean within-category similarity for each participant compared to the across-category similarity value for that category.

Overlap with VGG-16

We next assessed whether verbal recall reflects different information about what is encoded from VGG-16 by correlating each participant's similarity values with VGG-16 similarity values. We then applied a one-sample t-test to determine whether these correlations are significantly above zero.

Shared representational structure in recall

We next compared across-participant representational similarity by computing Pearson's correlation of within-category similarity matrices across participants and compared this 'true' across-participant similarity to a distribution of 10,000 permutations where pairwise image-image labels within each category were shuffled prior to correlation. Specifically, we calculated the correlation between all pairs of participants that were assigned to the same subset of categories

(e.g., indoor image group). Correlations were computed separately for each category and Fisher Z transformed before being averaged across categories and participants. Permutations were conducted at the group-level. 'True' across- participant representational similarity values were reverse transformed for reporting as Pearson's correlation values.

fMRI acquisition

Scanning was conducted on a Siemens 3T Prisma fMRI scanner with a 64-channel head coil at the Robert and Beverly Lewis Center for Neuroimaging at the University of Oregon. During the first scan day, high-resolution T2-weighted anatomical images were obtained in a ~15min session (voxel size 0.4 × 0.4 × 1.8mm; TR 11250.0ms; TE 84ms; in-plane acceleration factor (iPAT) 2). These scans were used to facilitate segmentation of hippocampal subfields. On the second day, participants underwent both anatomical and functional imaging. Before functional imaging, a whole-brain high-resolution anatomical image was collected for each participant using an MPRAGE sequence (isotropic voxel size 1mm; TR 2500.0ms; TE 3.43ms; TI 1100ms; flip angle 7°; in-plane acceleration factor (iPAT) 2). Partial brain functional images (excluding some prefrontal cortex) were collected using a T2*-weighted multi-band accelerated EPI sequence (isotropic voxel size 1.7mm; 80 axial slices; dist. factor 0%; field of view read 211mm; phase encoding direction A-P; base resolution 124; TR 1500ms; TE 33ms; flip angle 55°; echo spacing 0.6ms; phase partial Fourier 6/8; multiband acceleration factor 5). Each functional run consisted of 375 volumes and included 7.5s of lead-in and 7.5s of lead-out time at the beginning and end of each scan. EPI fieldmaps were acquired prior to functional runs (isotropic voxel size 1.7mm; TR 7550ms; TE 43.00ms; flip angle 90°; partial Fourier 6/8).

fMRI preprocessing

MRI data was first converted to Brain Imaging Data Structure (BIDS) using dcm2bids v3.1.1 (Boré et al., 2023) based on BIDS v1.8.0 (K. J. Gorgolewski et al., 2016). Next, functional and anatomical preprocessing was performed using fMRIPrep version 23.1.4 (Esteban et al., 2019) , a Nipype (K. Gorgolewski et al., 2011; K. J. Gorgolewski et al., 2017) based tool. The T1w image was corrected for intensity nonuniformity (INU) with N4BiasFieldCorrection v2.1.0 (Tustison et al., 2010) and skull-stripped using antsBrainExtraction.sh v2.1.0 (using the OASIS template). Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.1 (Dale et al.,

1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (Klein et al., 2017). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009) was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 (Avants et al., 2008) using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL v5.0.9; Zhang et al., 2001).

Functional data was motion corrected using mcflirt (FSL v5.0.9; Jenkinson et al., 2002). This was followed by co-registration to the corresponding T1w using boundary-based registration (Greve & Fischl, 2009) with 6 degrees of freedom, using bbregister (FreeSurfer v6.0.1). Motion correcting transformations, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation.

Physiological noise regressors were extracted applying CompCor (Behzadi et al., 2007). Principal components were estimated for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). A mask to exclude signal with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. Six tCompCor components were then calculated including only the top 5% variable voxels within that subcortical mask. For aCompCor, six components were calculated within the intersection of the subcortical mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run. Frame-wise displacement (Power et al., 2014) was calculated for each functional run using the implementation of Nipype.

Many internal operations of fMRIPrep use Nilearn (Abraham et al., 2014), principally within the BOLD-processing workflow. For more details of the pipeline see: https://fmriprep.readthedocs.io/en/20.2.0/workflows.html.

Copyright waiver

The above boilerplate text was automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts unchanged. It is released under the CC0 license.

We applied a threshold that no more than 15% of repetition times (TRs) could exceed a framewise displacement (FD) of 0.5mm within each run. Participants with more than one excluded functional run were excluded from analysis completely (n=2 excluded). For the included 38 participants, n=3 participants possessed singular runs that were beyond the threshold and excluded from analysis.

Brain masks generated by fMRIPrep from the T1 anatomical image were used to perform brain extraction for each of the six functional scans. For each functional run, each voxel was smoothed by 1.7mm FWHM Gaussian kernel and scaled between values 0-200 using AFNI (Cox, 1996; Cox & Hyde, 1997). All first level Generalized Linear Models (GLMs) were performed in participants' T1-aligned native space with AFNI. GLMs were calculated using the least-squares unitary (LSU) method with 3dREMLift (Abdulrahman & Henson, 2016). For each GLM, there was one regressor of interest representing a single scene across its two repetitions per scan, separately for encoding and retrieval phases within the scan. Other encoding trials and other retrieval trials were modeled as separate regressors, along with FD, xyz translation, and xyz rotation modeled as nuisance regressors. This model resulted in 36 beta-maps per scan (one map per scene/phase) which were converted to t-maps that represented the pattern of activity elicited by each scene/phase for each scan. Only t-maps corresponding to the encoding phase were used in analysis.

Regions of interest

We identify regions of interest (ROIs) using the HCP-MMP1 atlas (Fig. 4.2a; Glasser et al., 2016). First, we transform the atlas into each participant's T1-aligned native space. Next, we extract masks for two visual ROIs (early visual cortex and ventral occipitotemporal cortex) and two parietal ROIs (angular gyrus and ventral intraparietal sulcus) for each participant.

To extract hippocampal ROIs, we used the Automatic Segmentation of Hippocampal Subfields (ASHS) 64 toolbox (Yushkevich et al., 2015) with the upenn2017 atlas. This generated subfield ROIs in each participant's hippocampal body, including CA23DG (which included CA2, CA3, and dentate gyrus) and CA1 (Fig. 4.2b). The most anterior and posterior slices of the hippocampal body were manually determined for each participant based on the T2-weighted anatomical image. Each participant's subfield segmentations were also manually inspected to ensure the accuracy of the segmentation protocol. Then, each subfield ROI was transformed from

the T2 space into the T1 space using greedy (Yushkevich et al., 2016), with 0.2 voxel smoothing. All ROIs were again visually inspected following the transformation to T1 space to ensure the ROIs were anatomically correct.

For one participant, the T1 and T2 images (acquired in separate sessions) were not aligned. Therefore, we first resampled the T1 image with a T2 zoom. We then transformed the T2 image into the T1 space with antsApplyTransforms (Avants et al., 2008) with Lanczos resampling using the resampled T1 image as a reference and a transform file defined in ITK (McCormick et al., 2014; Yoo et al., 2002).

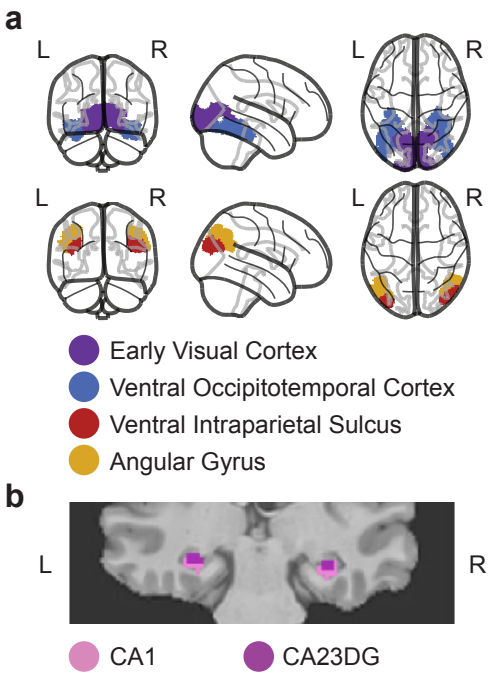

Figure 4.2. Regions of interest. a) Cortical ROIs. **b)** Hippocampal ROIs.

Analysis of neural data

Representational similarity analysis

As with recall and VGG-16, we generated similarity matrices using cosine similarity for neural activation t-maps. Neural similarity was calculated between pairs of different functional runs (e.g., run 1 t-map correlated with run 2 t-map). This resulted in separate neural matrices for each participant, ROI, and unique pair of runs (38 participants x 10 ROIs x 15 run pairs = 5700 neural similarity matrices). Next, a Fisher z-transformation was applied to each value in each

similarity matrix. We then subtracted mean across-category similarity from each within-category similarity value, separately for each category. Then, we extracted only within-category comparisons from each matrix. Finally, we average neural similarity matrices across all possible run pairs (38 participants x 10 ROIs = 380 neural similarity matrices).

Category sensitivity in neural activity

To determine which ROIs are sensitive to category-level information, we extracted within-category neural similarity values (which already have across-category similarity subtracted from them). For each ROI and subject, we computed the mean neural similarity value and conducted a one-sample t-test against zero for each region to determine which regions possess significantly higher within-category than across-category similarity.

Shared representational structure in neural activity

As with recall, we computed Pearson's correlation of within-category similarity matrices across participants and compared this 'true' across-participant similarity to a distribution of 10,000 permutations where pairwise image-image labels within each category were shuffled prior to correlation. Specifically, we calculated the correlation between all pairs of participants that were assigned to the same subset of categories (e.g., indoor image group). Correlations were computed separately for each category and ROI and Fisher Z transformed before being averaged across categories and participants. Permutations were conducted at the group-level. 'True' across-participant representational similarity values were reverse transformed for reporting as Pearson's correlation values.

Modeling neural representations of recall

To assess representations of within-subject verbal recall, we applied two different mixed effects models predicting neural similarity structure in each ROI as a function of similarity structure from other modalities (e.g., verbal recall and VGG-16). The first model involved predicting neural similarity structure as a function of VGG-16 similarity and participant-specific recall similarity. Our second model predicted neural similarity structure as a function of participant-specific recall similarity and normative recall similarity. Both models were

implemented using the lmerTest package in R with degrees of freedom calculated using Satterthwaite's approximation.

Prior to running the first model, we selected a single VGG-16 layer by applying a mixed effects model predicting within-category neural similarity from each VGG-16 layer with random effects of participant and category: neural ~ layer_1 + layer_3 + layer_5 + (1|participant) + (1|category). We found layer 3 had the largest beta values in four out of six of our ROIs and therefore chose layer 3 in our first model alongside participant-specific recall (Table 4.1). We then applied a mixed effects model separately for each ROI, predicting within-category values in the neural similarity matrices from within-category values in the VGG-16 layer 3 similarity matrix and recall similarity matrices, with random effects for participant and category: neural ~ layer_3 + participant-specific + (1|participant) + (1|category) (Table 4.2).

Next, we sought to disentangle the influence of shared recall from idiosyncratic recall and compare the relative influence of each in a model predicting neural similarity. Therefore, we ran a mixed effects model predicting neural similarity as a function of both normative recall and participant-specific recall with random intercepts for participant and category: neural ~ normative + participant-specific + (1|participant) + (1|category) (Table 4.3). We further computed the contrast between predictors using the glht() function from the multcomp package in R to determine the relative strength of each predictor in predicting neural similarity (Table 4.2; Hothorn et al., 2008).

Results

Behavioral measures of recall

Participants' vividness ratings significantly increased across the retrieval phases of the six learning rounds (mixed effects model: $t_{(188.10)} = 16.68$, $p < 0.001$), approaching ceiling (i.e., a rescaled value of 1.0) by the final round ($M = 0.96$, $SD = 0.07$). Accuracy on the post-scan associative memory test was an inclusion criterion for the experiment (no more than one error; see 'Methods'). Thus, the included sample of participants was able to subjectively recall the scene images with high vividness and to accurately remember the face-scene associations.

For the post-scan verbal recall test, the number of trials on which participants generated a response was an inclusion criterion (no more than one missing response). For trials on which participants responded, the mean number of words produced was 23.81 ($SD = 10.30$). There was no significant difference in the number of words recalled per scene category for participants that

studied 'indoor' categories (repeated measures ANOVA: $F_{(2,54)} = 0.28$, $p = 0.76$) or 'outdoor' categories (repeated measures ANOVA: $F_{(2,54)} = 1.02$, $p = 0.37$), or between the 'indoor' and 'outdoor' groups (Welch's $t_{(25.16)} = 1.73$, $p = 0.10$).

We next performed several analyses to validate text embeddings as a measure of verbal recall content (see Methods; Fig. 4.3a). As a first step, we tested whether recall embeddings contained category-level representational structure (e.g., that recall embeddings among 'pools' were more similar than recall embeddings between 'pools' and 'libraries'). Qualitatively, category-level representational structure can be readily perceived in a multidimensional scaling (MDS) plot, based on Euclidean distance, that visualizes the mean recall embeddings for all 36 scene stimuli (6 categories x 6 scenes per category; Fig. 4.3b). To statistically confirm category-level structure, we generated cosine similarity matrices based on each participant's recall (a sample participant's matrix is shown in Fig. 4.3c) and then compared within-category vs. across-category similarity using paired-sample t-tests. For each of the six scene categories, within-category similarity was significantly higher than across-category similarity (Pool: $t_{(18)} = 39.13$, $p < 0.001$; Library: $t_{(18)} = 23.60$, $p < 0.001$; Skating: $t_{(18)} = 30.98$, $p < 0.001$; Pond: $t_{(18)} = 52.22$, $p < 0.001$; Downtown: $t_{(18)} = 33.04$, $p < 0.001$; Stadium: $t_{(18)} = 22.91$, $p < 0.001$; Fig. 4.3d). These data clearly indicate that text embeddings reflected the high-level category structure of the recalled scenes.

As a second validation step, we tested whether within-category representational structure (e.g., the relative similarity with which a pair of 'library' scenes was recalled) was consistent (or shared) across participants. Consistency across participants would provide evidence that recall embeddings captured meaningful (non-random) within-category structure. To test this, we computed the Pearson's correlation of within-category similarity matrices across participants and compared this 'true' across-participant similarity to a distribution of 10,000 permutations where pairwise image-image labels within each category were shuffled prior to correlation. Specifically, we calculated the correlation between all pairs of participants that were assigned to the same subset of categories (e.g., indoor image group). Correlations were computed separately for each category and Fisher Z transformed before being averaged across categories and participants. Permutations were conducted at the group-level. The actual (reverse transformed) mean across-participant similarity value was only $r = 0.16$ but higher than each of the 10,000 permuted means. These data indicate that, within individual scene categories, there was a strong degree of alignment, across participants, in the representational structure of recall. Importantly, however, *differences* between

participants are, on their own, difficult to interpret. Differences could either reflect idiosyncratic details in participants' memories, or they could simply reflect measurement noise. We return to this point when we consider the fMRI data.

As a final validation step, we compared the within-category representational structure of recall embeddings to a 'ground truth' measure of scene content. Specifically, we selected three layers from the vision model VGG-16 (see Methods) that were intended to capture early (layer 1 max pooling), intermediate (layer 3 max pooling), and late (layer 5 max pooling) visual representations. As with the text embeddings, we used feature values from these layers to compute within-category representational structure (cosine similarity matrices). For each participant and each scene category, we then correlated the recall similarity matrix with each of the three VGG-16 similarity matrices. For each layer, we found significantly positive correlations (Layer 1: $t_{(37)} = 4.19$, $p < 0.001$; Layer 3: $t_{(37)} = 7.62$, $p < 0.001$; Layer 5: $t_{(37)} = 5.68$, $p < 0.001$). These data further confirm that within-category representational structure of recall embeddings captured meaningful information. However, as with the across-participant similarity analyses, differences between recall structure and VGG-16 structure are difficult to interpret as they could reflect meaningful idiosyncratic details of how participants remembered the scenes or they could reflect measurement noise.

fMRI measures of representational structure during encoding

All fMRI analyses focused on activity patterns during the study rounds (encoding) of the learning task, which preceded the post-scan verbal recall task. Pattern similarity was computed as the mean z-transformed cosine similarity between trial-specific beta values. Correlations were always restricted to trials from independent scan runs (see Methods). As with our analyses of verbal recall embeddings, we first tested for category-level representational structure. Specifically, for each of our ROIs, we computed the mean pattern similarity for stimuli from the same scene category (e.g., "library 1" and "library 2"; *within-category similarity*) and compared this against the mean pattern similarity for stimuli from different categories (e.g., "library 1" and "pool 2"; *across-category similarity*). Within-category similarity was significantly higher than across-category similarity for the visual ROIs (EVC: $t_{(37)} = 13.13$, $p < 0.001$; VOTC: $t_{(37)} = 10.11$, $p < 0.001$) and the parietal ROIs (vIPS: $t_{(37)} = 6.66$, $p < 0.001$; AG: $t_{(37)} = 4.48$, $p < 0.001$), but not for the hippocampal ROIs (CA1: $t_{(37)} = 0.33$, $p = 0.75$; CA23DG: $t_{(37)} = 13.13$, $p = 0.46$).

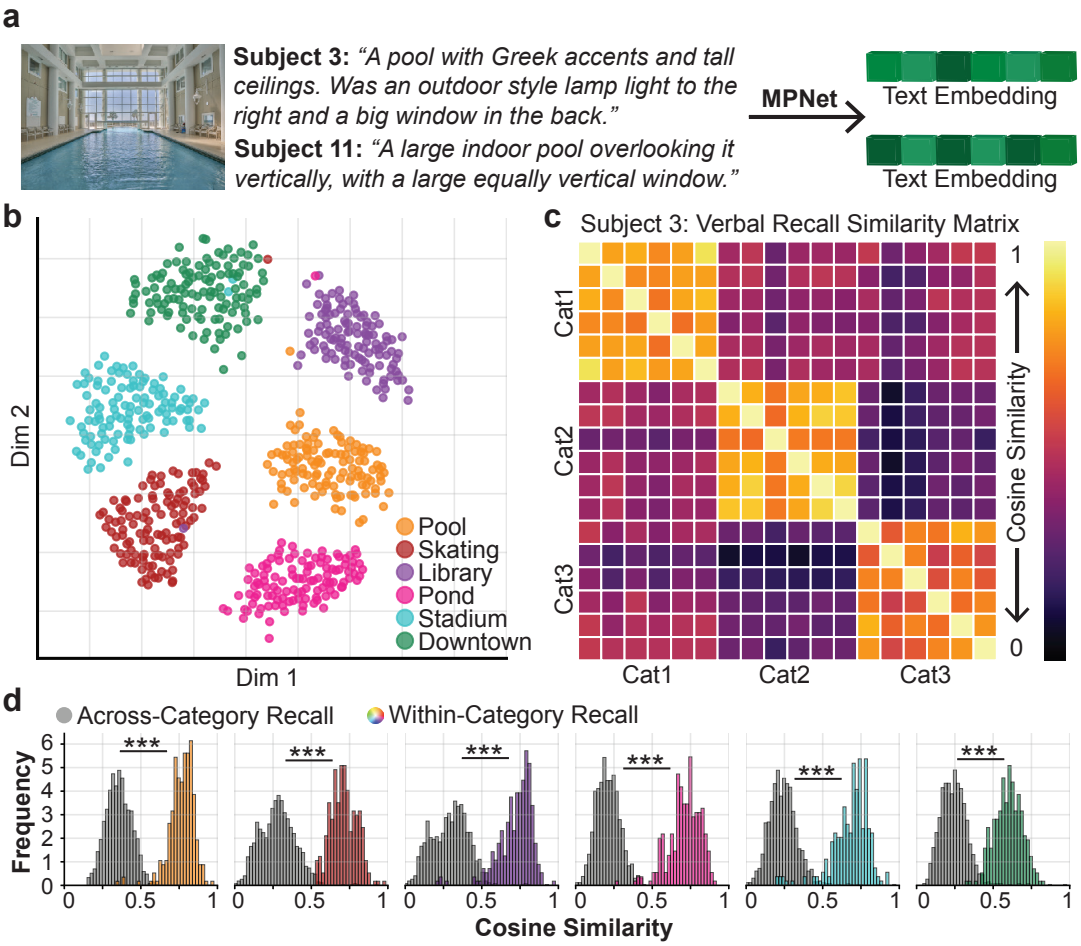

Figure 4.3. Verbal recall results. a) Verbal recall. Sample verbal recall descriptions from two participants. Recall descriptions were transformed into text embeddings using large language model MPNet. **b)** MDS plots of text embeddings from all participants, colored by image category. **c)** Recall similarity matrix. Sample recall cosine similarity matrix from a single participant. **d)** Within- VS across-category similarity. Distributions of within-category and across-category cosine similarity.

As with the verbal recall embeddings, we also tested whether fMRI measures of within-category representational structure were consistent (or shared) across participants. For each ROI, we correlated within-category similarity matrices between participants and compared this 'true' across-subject consistency to a distribution of N=10,000 permutations where pairwise image-image labels within each category were shuffled before correlation with other participants. Specifically, we calculated the correlation between all pairs of participants that were assigned to the same subset of categories (e.g., indoor image group). Correlations were computed separately for each category and Fisher Z transformed before being averaged across categories and participants. Permutations were conducted at the group-level. The actual (reverse transformed)

across-participant similarity was higher than most of the 10,000 permuted means for the visual ROIs (EVC: $r = 0.013$, > 10000 permutations; VOTC: $r = 0.019$, > 10000 permutations) and the parietal ROIs (vIPS: $r = 0.009$, > 10000 permutations; AG: $r = 0.0009$, > 9998 permutations), but not for the hippocampal ROIs (CA1: $r = 0.004$, > 9395 permutations; CA23DG: $r = -0.00007$, > 3963 permutations). Thus, for the visual and parietal ROIs there was modest, but statistically significant, alignment of within-category structure across participants. As with the verbal recall embeddings, 'misalignment' across participants is difficult to interpret as it could either reflect measurement noise or idiosyncratic differences in how participants encoded the scenes within each category.

Relating fMRI measures of encoding structure to text-based measures of recall structure

Of central interest, we next tested whether fMRI measures of representational structure during encoding were related to text embedding measures of representational structure during subsequent verbal recall. We were specifically interested in the degree to which a given participant's encoding structure was uniquely predicted by *that participant's* recall structure. We tested this using two different sets of mixed effects models. The first set of models included a 'ground truth' measure of representational structure defined by VGG-16. This allowed us to test whether verbal recall explained fMRI measures of representational structure *beyond* what was explained by VGG-16. The second set of models included a measure of normative recall similarity defined as the average representational structure of *other subjects'* verbal recall. This allowed us to test whether fMRI measures of representational structure captured participant-specific (idiosyncratic) recall content.

For the models that included VGG-16 measures of representational structure, we first identified the layer of VGG-16 that best predicted fMRI measures of representational structure. To do this, we ran mixed effects models (one for each ROI) in which within-category fMRI pattern similarity was the dependent measure and three VGG-16 layers were included as predictors (layer 1 max pooling, layer 3 max pooling, layer 5 max pooling). Qualitatively, the strongest relationship was observed for layer 3 (Table 4.1)—particularly for the visual and parietal ROIs. Having identified layer 3 as the 'best' predictor of fMRI representational structure, we then ran the critical mixed effects models (again, one for each ROI) that added participant-specific verbal recall as an *additional predictor* (alongside VGG-16 layer 3; Table 4.2). Note: for this model, significant

effects of layer 3 are shown in Fig. 4.4a but are not described here because this layer was specifically included in the model based on the results from the preceding, non-independent analysis.

Model 1 Statistics																		
	Layer 1 VGG-16				Layer 3 VGG-16				Layer 5 VGG-16				Layer 1-3		Layer 1-5		Layer 3-5	
ROI	β	SE	t-stat	p-val	β	SE	t-stat	p-val	β	SE	t-stat	p-val	z-stat	p-val	z-stat	p-val	z-stat	p-val
EVC	−0.02	0.01	−1.93	0.05	0.06	0.01	4.92	**<0.001**	0.02	0.01	1.71	0.09	3.95	**<0.001**	2.38	**0.02**	−1.75	0.08
VOTC	−0.01	0.01	−1.11	0.27	0.04	0.009	4.06	**<0.001**	−0.02	0.01	−2.39	**0.02**	2.97	**0.003**	−0.81	0.42	−3.74	**<0.001**
AG	0.006	0.01	0.53	0.59	0.03	0.01	2.49	**0.01**	−0.02	0.01	−1.65	0.10	1.10	0.27	−1.42	0.15	−2.42	**0.02**
vIPS	−0.03	0.02	−1.51	0.13	0.05	0.02	2.79	**0.005**	−0.03	0.02	−1.91	0.06	2.50	**0.01**	−0.24	0.81	−2.74	**0.006**
CA1	0.004	0.004	1.07	0.29	−0.004	0.004	−1.02	0.31	0.008	0.004	2.01	**0.04**	−1.28	0.20	0.63	0.53	1.84	0.07
DGCA23	0.006	0.006	1.00	0.32	0.004	0.006	0.60	0.55	−0.006	0.006	−1.01	0.32	−0.25	0.80	−1.32	0.19	−0.96	0.34

Table 4.1. Statistics from Model 1 (three VGG-16 layer predictors). For each ROI, beta values, standard error, t-statistics, and p-values are shown for each predictor and differences between predictors.

For each of the visual ROIs, we observed positive effects of participant-specific recall similarity on fMRI pattern similarity (EVC: ß = 0.03, SE = 0.009, $t_{(2738.13)} = 3.18$, $p = 0.002$; VOTC: ß = 0.03, SE = 0.007, $t_{(3179.26)} = 3.47$, $p < 0.001$). This indicates that when scenes from the same category had more similar activity patterns in visual cortex during encoding, they were associated with more similar subsequent verbal recall. Effects of participant-specific recall similarity were not significant in the parietal ROIs (vIPS: ß = 0.02, SE = 0.01, $t_{(3091.81)} = 1.65$, $p = 0.10$; AG: ß = -0.003, SE = 0.008, $t_{(1874.38)} = -0.32$, $p = 0.75$) or in CA1 (ß = -0.004, SE = 0.004, $t_{(244.52)} = -1.05$, $p = 0.30$). Interestingly, however, there was a negative effect in CA23DG (ß = -0.01, SE = 0.005, $t_{(439.43)} = -2.06$, $p = 0.04$). This negative effect indicates that CA23DG pattern similarity was *lower* when participant-specific recall similarity was *higher*. Taken together, results from this first set of models provide evidence that fMRI measures of encoding structure were related to the structure of subsequent verbal recall, even when we specifically accounted for the VGG-16 layer that best predicted fMRI pattern similarity (a 'ground truth' measure of similarity).

Model 2 Statistics										
	Layer 3 VGG-16				Participant-Specific Recall				VGG-Recall	
ROI	β	SE	t-stat	p-val	β	SE	t-stat	p-val	z-stat	p-val
EVC	0.05	0.008	6.12	**<0.001**	0.03	0.009	3.18	**0.002**	−1.79	0.07
VOTC	0.01	0.007	2.13	**0.03**	0.03	0.007	3.47	**<0.001**	0.98	0.33
AG	0.02	0.008	2.77	**0.006**	−0.003	0.008	−0.32	0.75	−1.99	**0.05**
vIPS	0.01	0.01	0.94	0.35	0.02	0.01	1.65	0.10	0.51	0.61
CA1	0.001	0.004	0.38	0.70	−0.004	0.004	−1.05	0.30	−0.95	0.34
DGCA23	0.006	0.005	1.25	0.21	−0.01	0.005	−2.06	**0.04**	−2.19	**0.03**

Table 4.2. Statistics from Model 2 (participant-specific recall and VGG-16 layer 3 predictors). For each ROI, beta values, standard error, t-statistics, and p-values are shown for each predictor and differences between predictors.

The second set of models was identical to the preceding models except that instead of including VGG-16 layer 3 max pooling, we included a normative measure of verbal recall similarity (i.e., mean recall similarity based on averaged text embeddings across all *other* participants in the experiment that studied the same set of scene images). Additionally, for these models we directly contrasted the two predictors (participant-specific verbal recall vs. normative verbal recall; Table 4.3).

For each of the visual ROIs, we again found positive effects of participant-specific recall similarity (EVC: ß = 0.04, SE = 0.009, $t_{(2819.86)} = 4.22$, $p < 0.001$; VOTC: ß = 0.03, SE = 0.007, $t_{(3200.30)} = 3.98$, $p < 0.001$), with no effect of normative recall similarity (EVC: ß = -0.01, SE = 0.009, $t_{(3313.08)} = -1.09$, $p = 0.28$; VOTC: ß = 0.01, SE = 0.008, $t_{(3318.38)} = 1.82$, $p = 0.07$). For EVC, but not VOTC, the participant-specific effect was significantly stronger than the normative effect (EVC: ß = 0.05, SE = 0.01, z = 3.75, $p < 0.001$; VOTC: ß = 0.01, SE = 0.01, z = 1.42, $p = 0.16$).

Among the parietal ROIs, there were positive effects of participant-specific recall similarity and normative recall similarity in vIPS (participant specific: ß = 0.03, SE = 0.01, $t_{(3155.50)} = 2.11$, $p = 0.04$; normative: ß = 0.07, SE = 0.01, $t_{(3055.66)} = 4.79$, $p < 0.001$), but no effect of either in AG (participant-specific: ß = 0.001, SE = 0.008, $t_{(2159.42)} = 0.15$, $p = 0.88$; normative: ß = -0.0008, SE = 0.008, $t_{(1199.98)} = -0.10$, $p = 0.92$). For vIPS, but not AG, the effect of normative recall similarity was significantly stronger than the effect of participant-specific recall similarity (vIPS: ß = -0.04, SE = 0.02, z = -2.12, $p = 0.03$; AG: ß = 0.002, SE = 0.01, z = 0.18, $p = 0.86$). It is important to note that the higher strength for normative recall may be attributed to an averaging confound as our normative recall measure was computed as an average across subjects. We

completed a supplementary permutation analysis to account for this difference in which we ran N=1000 permutations where a single other participant's recall was used in the model rather than averaged normative recall. Results from this analysis reveal comparable predictive ability of participant-specific and across-participant recall in vIPS (Fig. 4.5).

For CA1, there was no effect for participant-specific similarity or normative similarity (participant-specific: ß = -0.003, SE = 0.004, $t_{(710.55)}$ = -0.78, p = 0.44; normative: CA1: ß = 0.005, SE = 0.004, $t_{(204.25)}$ = 1.31, p = 0.19) and no difference between the two (ß = -0.009, SE = 0.005, z = -1.62, p = 0.10). For CA23DG, the effect of participant-specific recall similarity was numerically negative (as in the preceding analysis) but not significant, while the effect of normative recall similarity was positive and significant (participant-specific: ß = -0.007, SE = 0.005, $t_{(1097.06)}$ = -1.49, p = 0.14; normative: ß = 0.01, SE = 0.005, $t_{(312.19)}$ = 2.49, p = 0.01). The difference between these effects was also significant (CA23DG: ß = -0.02, SE = 0.007, z = -3.05, p = 0.002), indicating that the participant-specific effect in CA23DG was markedly lower (and in the opposite direction) relative to the positive effect of normative recall similarity.

Model 3 Statistics										
	Participant-Specific Recall				Normative Recall				Normative-Specific Recall	
ROI	β	SE	t-stat	p-val	β	SE	t-stat	p-val	z-stat	p-val
EVC	0.04	0.009	4.22	**<0.001**	−0.01	0.009	−1.09	0.28	3.75	**<0.001**
VOTC	0.03	0.007	3.98	**<0.001**	0.01	0.008	1.82	0.07	1.42	0.16
AG	0.001	0.008	0.15	0.88	−0.0008	0.008	−0.10	0.92	0.18	0.86
vIPS	0.03	0.01	2.11	**0.04**	0.07	0.01	4.79	**<0.001**	−2.12	**0.03**
CA1	−0.003	0.004	−0.78	0.44	0.005	0.004	1.31	0.19	−1.62	0.10
DGCA23	−0.007	0.005	−1.49	0.14	0.01	0.005	2.49	**0.01**	−3.05	**0.002**

Table 4.3. Statistics from Model 3 (participant-specific recall and normative recall predictors). For each ROI, beta values, standard error, t-statistics, and p-values are shown for each predictor and differences between predictors.

Taken together, results from the second set of models demonstrate that participant-specific recall similarity was related to fMRI measures of encoding structure, even when accounting for the normative structure of recall similarity from independent participants. However, the relative effects of participant-specific vs. normative recall similarity strongly differed across visual, parietal and hippocampal ROIs.

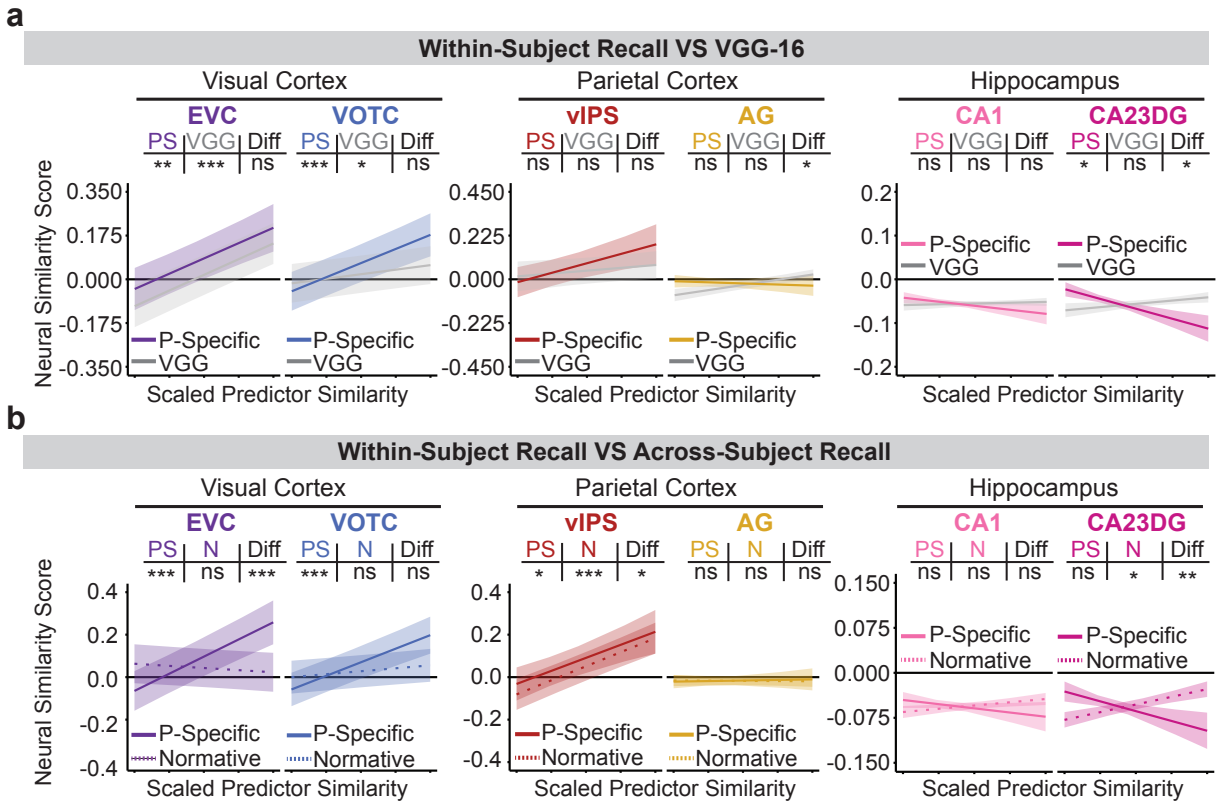

Figure 4.4. Neural mixed effects model results. a) Participant-specific VS VGG-16 model. For each ROI, regression lines from mixed effects model beta values shown. Colored lines reflect relationship between neural similarity and participant-specific similarity, and gray lines reflect relationship between neural similarity and VGG-16 similarity. Significance tables above indicate significance of each predictor and significant differences. **b)** Participant-specific VS normative model. For each ROI, regression lines from mixed effects model beta values shown. Solid lines reflect relationship between neural similarity and participant-specific similarity, and dashed lines reflect relationship between neural similarity and normative similarity. Significance tables above indicate significance of each predictor and significant differences.

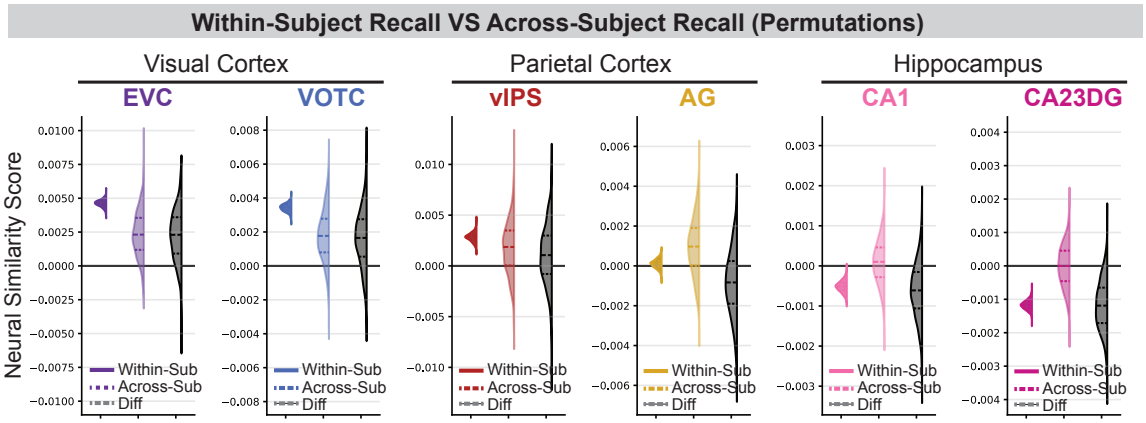

Figure 4.5. Supplementary permutation results. Participant-specific (i.e., within-sub) VS random across-subject models. For each ROI, distributions of beta values from N=1000 permutations of mixed effects model beta values shown. Solid colored distributions reflect relationship between neural similarity and participant-specific similarity, faded colored distributions reflect relationship between neural similarity and across-participant similarity, and gray distributions reflect difference between the two distributions.

General Discussion

Here, we assessed the neural representational structure of naturalistic scene images during encoding, and its relationship with subsequent memory content. Specifically, we used verbal recall as our measure of subsequent memory content and separately defined participant-specific and normative components of recall. Next, we determined the relationship between representational structure in visual, parietal, and hippocampal ROIs and representational structure in participant-specific and normative verbal recall. We first establish that verbal recall possesses unique information about neural representational structure that cannot be accounted for by an 'objective' measure like VGG-16. Next, we show that visual ROIs better reflect participant-specific structure in verbal recall over normative structure – a pattern which does not extend to parietal ROIs. Finally, we provide evidence that representational structure in hippocampal CA23DG is negatively associated with participant-specific memory structure but positively associated with normative memory structure. Taken together, these findings provide evidence for distinct and complementary roles of different ROIs in representing idiosyncratic and shared memory content.

Many fMRI studies have used verbal recall as a measure of subsequent memory. However, they are often limited to univariate measures of fMRI activity and recall of simple word lists (Brassen et al., 2006; H. Kim, 2011; Long et al., 2010; Staresina & Davachi, 2006; Strange et al., 2002; Wagner, 1998). More recent work has explored how neural representations during encoding

are related to subsequent verbal recall descriptions (Saalasti et al., 2019; Wang et al., 2023) and complex narratives (Lee & Chen, 2022; Masís-Obando et al., 2022; Nguyen et al., 2019). These studies present verbal recall descriptions as a rich, naturalistic, and flexible measurement of memory content – flexibility which could allow individuals to spontaneously express idiosyncrasies in their memories. While memory differences have been intentionally manipulated by modulating attention, context, etc. (Anderson & Pichert, 1978; De Soares et al., 2024; Uncapher & Rugg, 2009), spontaneous idiosyncratic differences are much harder to capture (Finn et al., 2020). For one, idiosyncratic differences are difficult to distinguish from noise (Finn et al., 2020). For another, it is difficult to isolate idiosyncratic memory from shared memory, leading many studies to define idiosyncratic memory as dissimilarity across participants (Nguyen et al., 2019), rather than explicitly defining separate measures of participant-specific and across-participant memory content. As such, in the current chapter we define separate predictors for participant-specific and across-participant verbal recall, and conceptualize idiosyncratic memory content as the difference between these measures.

In comparing the relative strength of participant-specific and normative structure in cortical ROIs, we find that visual and parietal cortex dissociably reflect participant-specific and/or normative recall structure. This difference validates that our measures of idiosyncratic and shared memory content do indeed reflect differences in what is encoded into memory (rather than simple individual differences in how participants communicate or chose to perform the task). Specifically, we find that visual cortex better reflects participant-specific memory content. The fact that the participant-specific effect was significantly stronger than the normative effect (for EVC) is particularly striking given that normative recall similarity was derived from many more participants. If, for example, each participant shared the same recall structure (plus random noise), then the normative measure would have a much higher signal-to-noise ratio than the participant-specific measure. The more idiosyncratic nature of visual cortex is consistent with prior literature indicating visual cortex as being highly influenced by fluctuations in attention (Chapman et al., 2023; Kok et al., 2013; Maunsell & Treue, 2006), potentially biasing visual cortex to possess more variable (and idiosyncratic) representations of visual stimuli. Meanwhile, we find that parietal cortex better reflects normative memory content. It is important to note that the higher representational strength may be attributed to an averaging bias as our normative recall measure was computed as an average across subjects. A permutation analysis accounting for this difference

reveals comparable predictive ability of participant-specific and normative recall in parietal cortex. Nevertheless, in comparison to visual cortex, parietal cortex possesses relatively stronger representations of normative memory content. This finding suggests parietal cortex better represents shared memory content. For example, some studies present parietal cortex as representing more abstract and generalizable concepts (Nakai et al., 2025; Yeshurun et al., 2021), which may lead to a preference for shared features in memory (Chen et al., 2017; Hasson et al., 2008; Nastase et al., 2019; Wang et al., 2023). While our results are potentially indicative of different functional roles in visual and parietal cortex, we must be cautious with generalizations, as it is likely that our findings depend on the nature of our stimuli. Here, we use highly similar scene stimuli which possess more low-level visual variability than semantic variability, potentially biasing idiosyncratic representations to visual cortex. In contrast, abstract stimuli (e.g., fractals), may possess the opposite effect: less visual variability but more interpretation variability (Bellemare-Pepin et al., 2022; Rogowitz & Voss, 1990; Taylor et al., 2017; Vessel & Rubin, 2010) may lead to stronger idiosyncratic representations in parietal cortex.

In addition to the differences between visual and parietal cortex, we observe an interesting pattern in hippocampal subregion CA23DG where CA23DG negatively reflects participant-specific memory structure. This pattern is most strongly evident in our model predicting CA23DG representational structure from participant-specific verbal recall and VGG-16 structure. Even in the model combining participant-specific and normative recall, the relationship between CA23DG and participant-specific recall is numerically negative (lower strength of effect likely due to overlapping features in participant-specific and normative recall), and significantly lower than its relationship with normative recall. These findings, while counterintuitive, may be consistent with prior work implicating CA23DG in 'hippocampal differentiation' (Chanales et al., 2017; Dimsdale-Zucker et al., 2018; Favila et al., 2016; Hulbert & Norman, 2015; G. Kim et al., 2017; Mirjalili et al., 2025; Schlichting et al., 2015; Wanjia et al., 2021; Wing et al., 2020), wherein highly overlapping events are adaptively separated to be represented more *dissimilarly* to prevent memory interference. While studies have indicated that hippocampal differentiation is associated with higher memory success (Favila et al., 2016; Hulbert & Norman, 2015; Wanjia et al., 2021; Wing et al., 2020), it remains unclear whether this separation in neural representations extends to a separation in behavioral measures of memory content (Shao et al., 2023). Based on our findings, it is possible that CA23DG may be resolving memory interference based on subjective

(participant-specific) assessments of stimulus overlap. Moreover, this repulsion may occur so that memory content itself need not to be distorted. However, these findings are preliminary, and further work is needed to better understand the interaction between hippocampal differentiation and memory content.

In conclusion, this chapter provides evidence for (1) dissociable and complementary roles of visual and parietal cortex in representing idiosyncratic and shared memory content, and (2) adaptive separation of hippocampal representations based on idiosyncratic assessments of overlap. Further, by comparing representational structure across different measurements (i.e., neural activity, verbal recall, VGG-16), and within- and across-participants, we gain important insight into the nature and generalizability of idiosyncratic and shared representations of visual stimuli. Moreover, this work provides insight into how idiosyncratic differences during encoding can shape how we experience and remember the world, which may vary in clinical and neurodivergent populations (Hahamy et al., 2015; Hasson et al., 2009; Lin et al., 2026; Mäntylä & Bäckman, 1990; Matsumoto et al., 2023; Perl et al., 2023).

References Cited

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., … Zheng, X. (2016). *TensorFlow: A system for large-scale machine learning* (arXiv:1605.08695). arXiv. https://doi.org/10.48550/arXiv.1605.08695

Abdulrahman, H., & Henson, R. N. (2016). Effect of trial-to-trial variability on optimal event-related fMRI design: Implications for Beta-series correlation and multi-voxel pattern analysis. *NeuroImage*, *125*, 756–766. https://doi.org/10.1016/j.neuroimage.2015.11.009

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/fninf.2014.00014

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, *17*(1), 1–12. https://doi.org/10.1016/S0022-5371(78)90485-1

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology* (pp. xix, 317). Cambridge University Press.

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Bellemare-Pepin, A., Harel, Y., O'Byrne, J., Mageau, G., Dietrich, A., & Jerbi, K. (2022). Processing visual ambiguity in fractal patterns: Pareidolia as a sign of creativity. *iScience*, *25*(10), 105103. https://doi.org/10.1016/j.isci.2022.105103

Boré, A., Guay, S., Bedetti, C., Meisler, S., & GuenTher, N. (2023). *Dcm2Bids* (Version 3.1.1) [Computer software]. https://doi.org/10.5281/zenodo.8436509

Brassen, S., Weber-Fahr, W., Sommer, T., Lehmbeck, J. T., & Braus, D. F. (2006). Hippocampal–prefrontal encoding activation predicts whether words can be successfully recalled or only recognized. *Behavioural Brain Research*, *171*(2), 271–278. https://doi.org/10.1016/j.bbr.2006.04.002

Chanales, A. J. H., Oza, A., Favila, S. E., & Kuhl, B. A. (2017). Overlap among Spatial Memories Triggers Repulsion of Hippocampal Representations. *Current Biology*, *27*(15), 2307-2317.e5. https://doi.org/10.1016/j.cub.2017.06.057

Chapman, A. F., Chunharas, C., & Störmer, V. S. (2023). Feature-based attention warps the perception of visual features. *Scientific Reports*, *13*(1), 6487. https://doi.org/10.1038/s41598-023-33488-2

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Chollet, F. (2015). *Keras* [Computer software]. https://github.com/fchollet/keras

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014

Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, *10*(4–5), 171–178. https://doi.org/10.1002/(sici)1099-1492(199706/08)10:4/5%3C171::aid-nbm453%3E3.0.co;2-l

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

De Soares, A., Kim, T., Mugisho, F., Zhu, E., Lin, A., Zheng, C., & Baldassano, C. (2024). Top-down attention shifts behavioral and neural event boundaries in narratives with overlapping event scripts. *Current Biology*, *34*(20), 4729-4742.e5. https://doi.org/10.1016/j.cub.2024.09.013

Dimsdale-Zucker, H. R., Ritchey, M., Ekstrom, A. D., Yonelinas, A. P., & Ranganath, C. (2018). CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nature Communications*, *9*(1), 294. https://doi.org/10.1038/s41467-017-02752-1

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Favila, S. E., Chanales, A. J. H., & Kuhl, B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nature Communications*, *7*(1), 11066. https://doi.org/10.1038/ncomms11066

Finn, E. S., Glerean, E., Khojandi, A. Y., Nielson, D., Molfese, P. J., Handwerker, D. A., & Bandettini, P. A. (2020). Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *NeuroImage*, *215*, 116828. https://doi.org/10.1016/j.neuroimage.2020.116828

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage, Organization for Human Brain Mapping 2009 Annual Meeting*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Van Essen, D. C. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171–178. https://doi.org/10.1038/nature18933

Gordon, A. M., Rissman, J., Kiani, R., & Wagner, A. D. (2014). Cortical Reinstatement Mediates the Relationship Between Content-Specific Encoding Activity and Subsequent Recollection Decisions. *Cerebral Cortex*, *24*(12), 3350–3364. https://doi.org/10.1093/cercor/bht194

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, *5*, 13. https://doi.org/10.3389/fninf.2011.00013

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*(1), 160044. https://doi.org/10.1038/sdata.2016.44

Gorgolewski, K. J., Esteban, O., Gage Ellis, D., Notter, M. P., Ziegler, E., Johnson, H., Hamalainen, C., Yvernault, B., Burns, C., Manhães-Savio, A., Jarecka, D., Markiewicz, C. J., Salo, T., Clark, D., Waskom, M., Wong, J., Modat, M., Dewey, B. E., Clark, M. G., … Ghosh, S. (2017). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in Python. 0.13.1. *Zenodo*. https://doi.org/10.5281/zenodo.581704

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Hahamy, A., Behrmann, M., & Malach, R. (2015). The idiosyncratic brain: Distortion of spontaneous connectivity patterns in autism spectrum disorder. *Nature Neuroscience*, *18*(2), 302–309. https://doi.org/10.1038/nn.3919

Hasson, U., Avidan, G., Gelbard, H., Vallines, I., Harel, M., Minshew, N., & Behrmann, M. (2009). Shared and idiosyncratic cortical activation patterns in autism revealed under continuous real-life viewing conditions. *Autism Research: Official Journal of the International Society for Autism Research*, *2*(4), 220–231. https://doi.org/10.1002/aur.89

Hasson, U., Furman, O., Clark, D., Dudai, Y., & Davachi, L. (2008). Enhanced Intersubject Correlations during Movie Viewing Correlate with Successful Episodic Encoding. *Neuron*, *57*(3), 452–462. https://doi.org/10.1016/j.neuron.2007.12.009

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Heusser, A. C., Fitzpatrick, P. C., & Manning, J. R. (2021). Geometric models reveal behavioural and neural signatures of transforming experiences into memories. *Nature Human Behaviour*, *5*(7), 905–919. https://doi.org/10.1038/s41562-021-01051-6

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal. Biometrische Zeitschrift*, *50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Hulbert, J. C., & Norman, K. A. (2015). Neural Differentiation Tracks Improved Recall of Competing Memories Following Interleaved Study and Retrieval Practice. *Cerebral Cortex*, *25*(10), 3994–4008. https://doi.org/10.1093/cercor/bhu284

Jääskeläinen, I. P., Koskentalo, K., Balk, M. H., Autti, T., Kauramäki, J., Pomren, C., & Sams, M. (2008). Inter-Subject Synchronization of Prefrontal Cortex Hemodynamic Activity During Natural Viewing. *The Open Neuroimaging Journal*, *2*, 14–19. https://doi.org/10.2174/1874440000802010014

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1016/s1053-8119(02)91132-8

Ki, J. J., Kelly, S. P., & Parra, L. C. (2016). Attention Strongly Modulates Reliability of Neural Responses to Naturalistic Narrative Stimuli. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *36*(10), 3092–3101. https://doi.org/10.1523/JNEUROSCI.2942-15.2016

Kim, G., Norman, K. A., & Turk-Browne, N. B. (2017). Neural Differentiation of Incorrectly Predicted Memories. *The Journal of Neuroscience*, *37*(8), 2022–2031. https://doi.org/10.1523/JNEUROSCI.3272-16.2017

Kim, H. (2011). Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage*, *54*(3), 2446–2461. https://doi.org/10.1016/j.neuroimage.2010.09.045

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Neto, E. C., & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, *13*(2), e1005350. https://doi.org/10.1371/journal.pcbi.1005350

Kok, P., Brouwer, G. J., Gerven, M. A. J. van, & Lange, F. P. de. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *Journal of Neuroscience*, *33*(41), 16275–16284. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Lee, H., & Chen, J. (2022). Predicting memory from the network structure of naturalistic events. *Nature Communications*, *13*(1), 4235. https://doi.org/10.1038/s41467-022-31965-2

Leeuw, J. R. de, Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an Open-Source Collaborative Ecosystem of Behavioral Experiments. *Journal of Open Source Software*, *8*(85), 5351. https://doi.org/10.21105/joss.05351

Lin, H.-Y., Breakspear, M., & Mottron, L. (2026). From heterogeneity to idiosyncrasy in the autistic brain. *Nature Mental Health*, *4*(3), 346–359. https://doi.org/10.1038/s44220-026-00601-z

Long, N. M., Öztekin, I., & Badre, D. (2010). Separable Prefrontal Cortex Contributions to Free Recall. *The Journal of Neuroscience*, *30*(33), 10967–10976. https://doi.org/10.1523/JNEUROSCI.2611-10.2010

Mäntylä, T., & Bäckman, L. (1990). Encoding variability and age-related retrieval failures. *Psychology and Aging*, *5*(4), 545–550. https://doi.org/10.1037//0882-7974.5.4.545

Masís-Obando, R., Norman, K. A., & Baldassano, C. (2022). Schema representations in distinct brain networks support narrative memory during encoding and retrieval. *eLife*, *11*, e70445. https://doi.org/10.7554/eLife.70445

Matsumoto, Y., Nishida, S., Hayashi, R., Son, S., Murakami, A., Yoshikawa, N., Ito, H., Oishi, N., Masuda, N., Murai, T., Friston, K., Nishimoto, S., & Takahashi, H. (2023). Disorganization of Semantic Brain Networks in Schizophrenia Revealed by fMRI. *Schizophrenia Bulletin*, *49*(2), 498–506. https://doi.org/10.1093/schbul/sbac157

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. https://doi.org/10.1016/j.tins.2006.04.001

McCormick, M., Liu, X., Jomier, J., Marion, C., & Ibanez, L. (2014). ITK: Enabling reproducible research and open science. *Frontiers in Neuroinformatics*, *8*, 13. https://doi.org/10.3389/fninf.2014.00013

Mirjalili, S., Wanjia, G., Grätz, D., Wang, E., Mayr, U., & Kuhl, B. A. (2025). Hippocampal transformations occur along dimensions of memory interference. *bioRxiv*, 2025.10.13.682242. https://doi.org/10.1101/2025.10.13.682242

Nakai, T., Kubo, R., & Nishimoto, S. (2025). Cortical representational geometry of diverse tasks reveals subject-specific and subject-invariant cognitive structures. *Communications Biology*, *8*(1), 713. https://doi.org/10.1038/s42003-025-08134-4

Nastase, S. A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience*, *14*(6), 667–685. https://doi.org/10.1093/scan/nsz037

Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage*, *184*, 161–170. https://doi.org/10.1016/j.neuroimage.2018.09.010

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perl, O., Duek, O., Kulkarni, K. R., Gordon, C., Krystal, J. H., Levy, I., Harpaz-Rotem, I., & Schiller, D. (2023). Neural patterns differentiate traumatic from sad autobiographical memories in PTSD. *Nature Neuroscience*, *26*(12), 2226–2236. https://doi.org/10.1038/s41593-023-01483-5

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 10.1016/j.neuroimage.2013.08.048. https://doi.org/10.1016/j.neuroimage.2013.08.048

Rehurek, R. (2011). Gensim–python framework for vector space modelling. In *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic.* (Vol. 3, Issue 2). https://cir.nii.ac.jp/crid/1370869855565691916

Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks* (arXiv:1908.10084). arXiv. https://doi.org/10.48550/arXiv.1908.10084

Rogowitz, B. E., & Voss, R. (1990). Shape perception and low-dimension fractal boundary contours. *Human Vision and Electronic Imaging: Models, Methods, and Applications*, *1249*, 387–394. https://doi.org/10.1117/12.19691

Saalasti, S., Alho, J., Bar, M., Glerean, E., Honkela, T., Kauppila, M., Sams, M., & Jääskeläinen, I. P. (2019). Inferior parietal lobule and early visual areas support elicitation of individualized meanings during narrative listening. *Brain and Behavior*, *9*(5), e01288. https://doi.org/10.1002/brb3.1288

Sava-Segal, C., Richards, C., Leung, M., & Finn, E. S. (2023). Individual differences in neural event segmentation of continuous experiences. *Cerebral Cortex*, *33*(13), 8164–8178. https://doi.org/10.1093/cercor/bhad106

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*(1), 8151. https://doi.org/10.1038/ncomms9151

Shao, X., Li, A., Chen, C., Loftus, E. F., & Zhu, B. (2023). Cross-stage neural pattern similarity in the hippocampus predicts false memory derived from post-event inaccurate information. *Nature Communications*, *14*(1), 2299. https://doi.org/10.1038/s41467-023-38046-y

Sheng, J., Wang, S., Zhang, L., Liu, C., Shi, L., Zhou, Y., Hu, H., Chen, C., & Xue, G. (2023). Intersubject similarity in neural representations underlies shared episodic memory content. *Proceedings of the National Academy of Sciences*, *120*(35), e2308951120. https://doi.org/10.1073/pnas.2308951120

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv:1409.1556). arXiv. https://doi.org/10.48550/arXiv.1409.1556

Song, H., Finn, E. S., & Rosenberg, M. D. (2021). Neural signatures of attentional engagement during narratives and its consequences for event memory. *Proceedings of the National Academy of Sciences*, *118*(33), e2021905118. https://doi.org/10.1073/pnas.2021905118

Song, K., Tan, X., Qin, T., Lu, J., & Liu, T.-Y. (2019). *MASS: Masked Sequence to Sequence Pre-training for Language Generation*. https://doi.org/10.48550/ARXIV.1905.02450

Staresina, B. P., & Davachi, L. (2006). Differential encoding mechanisms for subsequent associative recognition and free recall. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(36), 9162–9172. https://doi.org/10.1523/JNEUROSCI.2877-06.2006

Staresina, B. P., Duncan, K. D., & Davachi, L. (2011). Perirhinal and Parahippocampal Cortices Differentially Contribute to Later Recollection of Object- and Scene-Related Event Details. *Journal of Neuroscience*, *31*(24), 8739–8747. https://doi.org/10.1523/JNEUROSCI.4978-10.2011

Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable Human Perirhinal, Hippocampal, and Parahippocampal Roles during Verbal Encoding. *The Journal of Neuroscience*, *22*(2), 523–528. https://doi.org/10.1523/JNEUROSCI.22-02-00523.2002

Taylor, R. P., Martin, T. P., Montgomery, R. D., Smith, J. H., Micolich, A. P., Boydston, C., Scannell, B. C., Fairbanks, M. S., & Spehar, B. (2017). Seeing shapes in seemingly random spatial patterns: Fractal analysis of Rorschach inkblots. *PLOS ONE*, *12*(2), e0171289. https://doi.org/10.1371/journal.pone.0171289

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Uncapher, M. R., & Rugg, M. D. (2009). Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. *The Journal of Neuroscience*, *29*(25), 8270–8279. https://doi.org/10.1523/JNEUROSCI.1043-09.2009

Vessel, E. A., & Rubin, N. (2010). Beauty and the beholder: Highly individual taste for abstract, but not real-world images. *Journal of Vision*, *10*(2), 18. https://doi.org/10.1167/10.2.18

Wagner, A. D. (1998). Building Memories: Remembering and Forgetting of Verbal Experiences as Predicted by Brain Activity. *Science*, *281*(5380), 1188–1191. https://doi.org/10.1126/science.281.5380.1188

Wagner, A. D., Koutstaal, W., & Schacter, D. L. (1999). When encodong yields remembering: Insights from event-related neuroimaging. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *354*(1387), 1307–1324. https://doi.org/10.1098/rstb.1999.0481

Wang, Y., Lee, H., & Kuhl, B. A. (2023). *Mapping multidimensional content representations to neural and behavioral expressions of episodic memory* [Preprint]. Neuroscience. https://doi.org/10.1101/2023.01.09.523096

Wanjia, G., Favila, S. E., Kim, G., Molitor, R. J., & Kuhl, B. A. (2021). Abrupt hippocampal remapping signals resolution of memory interference. *Nature Communications*, *12*(1), 4816. https://doi.org/10.1038/s41467-021-25126-0

Watanabe, T., Hirose, S., Wada, H., Katsura, M., Chikazoe, J., Jimura, K., Imai, Y., Machida, T., Shirouzu, I., Miyashita, Y., & Konishi, S. (2011). Prediction of subsequent recognition performance using brain activity in the medial temporal lobe. *NeuroImage*, *54*(4), 3085–3092. https://doi.org/10.1016/j.neuroimage.2010.10.066

Wing, E. A., Geib, B. R., Wang, W.-C., Monge, Z., Davis, S. W., & Cabeza, R. (2020). Cortical Overlap and Cortical-Hippocampal Interactions Predict Subsequent True and False Memory. *Journal of Neuroscience*, *40*(9), 1920–1930. https://doi.org/10.1523/JNEUROSCI.1766-19.2020

Yeshurun, Y., Nguyen, M., & Hasson, U. (2021). The default mode network: Where the idiosyncratic self meets the shared social world. *Nature Reviews. Neuroscience*, *22*(3), 181–192. https://doi.org/10.1038/s41583-020-00420-w

Yoo, T. S., Ackerman, M. J., Lorensen, W. E., Schroeder, W., Chalana, V., Aylward, S., Metaxas, D., & Whitaker, R. (2002). Engineering and algorithm design for an image processing Api: A technical report on ITK--the Insight Toolkit. *Studies in Health Technology and Informatics*, *85*, 586–592.

Yushkevich, P. A., Amaral, R. S. C., Augustinack, J. C., Bender, A. R., Bernstein, J. D., Boccardi, M., Bocchetta, M., Burggren, A. C., Carr, V. A., Chakravarty, M. M., Chételat, G., Daugherty, A. M., Davachi, L., Ding, S.-L., Ekstrom, A., Geerlings, M. I., Hassan, A., Huang, Y., Iglesias, J. E., … Hippocampal Subfields Group (HSG). (2015). Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: Towards a harmonized segmentation protocol. *NeuroImage*, *111*, 526–541. https://doi.org/10.1016/j.neuroimage.2015.01.004

Yushkevich, P. A., Pluta, J., Wang, H., Wisse, L. E. M., Das, S., & Wolk, D. (2016). IC-P-174: Fast Automatic Segmentation of Hippocampal Subfields and Medial Temporal Lobe Subregions In 3 Tesla and 7 Tesla T2-Weighted MRI. *Alzheimer's & Dementia*, *12*(7S_Part_2), P126–P127. https://doi.org/10.1016/j.jalz.2016.06.205

Zhang, Y., Brady, M., & Smith, S. A. (2001). Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the Expectation Maximization Algorithm. *IEEE Transactions on Medical Imaging*, *20*, 45–57. https://doi.org/10.1109/42.906424

CHAPTER V
GENERAL DISCUSSION

Ongoing experience is filled with highly overlapping events, which shape episodic memory. While extensive work has explored the effects of overlap on memory, they are largely limited to highly controlled stimuli and simple measures of memory outcomes (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024), limiting generalizability of results to more naturalistic and real-world contexts. As such, this dissertation sought to explore how overlap influences memory using more naturalistic stimuli and methods. Specifically, I utilized verbal recall as the primary measure of memory content in order to consider broader and more complex questions about episodic memory. The primary goals of this dissertation were to understand: (1) how overlapping naturalistic scene images affect verbal recall, and (2) how an individual's idiosyncrasies affect encoding of overlapping stimuli.

In **Chapter II**, we conducted a pair of experiments to assess how overlap influences verbal recall of naturalistic scene images. In a first experiment, we showed that overlap between multiple stimuli leads to separation in verbal recall, analogous with accounts of memory repulsion (Chanales et al., 2021; Drascher & Kuhl, 2022) and/or an increase in memory precision (Drascher & Kuhl, 2022). In the second experiment, we used pairs of overlapping stimuli and found no evidence of separation; rather, we observed an effect wherein the specific overlapping competitor influences verbal recall in a consistent but nondirectional manner. Taken together, these experiments provide insight into the subtle influences of overlap on verbal recall for naturalistic images.

In **Chapter III**, we took a closer look at the second experiment from **Chapter II**. Specifically, we developed a new analysis to consider whether verbal recall prioritizes diagnostic patches of a scene image. We found that patch importance is consistent among participants with the same competitor image. Moreover, we demonstrated that patch importance is driven by objective dissimilarity between the two overlapping scene images at that patch – suggesting verbal recall prioritizes patches of an image that possess diagnostic features. This work serves as a proof of concept for a new method, providing insight into the dynamic influences of overlap on verbal recall.

In **Chapter IV**, we conducted an fMRI study to assess how neural representational structure during encoding of overlapping scene images relates to idiosyncratic and shared

subsequent memory content. As in prior chapters, we used verbal recall as our measure of memory content, and dissociate participant-specific (idiosyncratic) features in recall from normative (shared) features in recall. We found that visual cortex better reflects participant-specific verbal recall while parietal cortex reflects normative verbal recall. Additionally, we showed that hippocampus CA23DG is negatively associated with participant-specific verbal recall – consistent with hippocampal differentiation (Chanales et al., 2017; Dimsdale-Zucker et al., 2018; Favila et al., 2016; Hulbert & Norman, 2015; Kim et al., 2017; Mirjalili et al., 2025; Schlichting et al., 2015; Wanjia et al., 2021; Wing et al., 2020). Taken together, this chapter provides evidence for distinct and complementary roles of visual cortex, parietal cortex, and hippocampus, in representing idiosyncratic and shared memory content for overlapping scene images.

Quantifying complex visual memories

Across all three empirical chapters, we used naturalistic scene images as our stimuli. This approach sets the current work apart from prior studies of visual memory in that our stimuli possess new and uncontrolled sources of variability. Prior research, spanning decades of work, primarily considered simple features for controlled stimuli (e.g., color, orientation, etc.; Bae & Luck, 2017; Chanales et al., 2021; Chapman et al., 2023; Chunharas et al., 2022; Golomb, 2015; Magnussen & Greenlee, 1999; Morrow & Clewett, 2025; Rademaker et al., 2015). In order to quantify memory for more complex visual stimuli, we used and developed new methods for parsing stimulus features.

In **Chapter III**, we developed a new method to zoom in on stimulus features that are prioritized in recall due to overlap. Typically, studies probing memory for stimulus features use highly controlled stimuli (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024). Here, we obtained a rich measure of memory (i.e., verbal recall), and mapped its features to different parts of an image. Specifically, we segmented images into smaller image patches and assessed how well verbal recall maps to these patches. This method opens new possibilities for quantifying memories of complex visual stimuli. Specifically, as long as there is some quantifiable measure (e.g., verbal recall text embeddings, gaze patterns, etc.), we can assess localized patch preferences by mapping these measures to those regions. In this manner, we may consider subtle effects in memory even with rich and high-dimensional measures of memory content.

Another unique method we used across all three empirical chapters is large language models to quantify verbal recall descriptions. The use of verbal recall to quantify visual memory could be considered counterintuitive; verbal recall descriptions are dominated by semantic information, which may not be sufficient to characterize memory for a visual stimulus (especially the highly overlapping visual stimuli used in this dissertation). However, verbal recall descriptions have recently gained popularity, and have been used to quantify memory for scene images (Wang et al., 2023) and movies (Chen et al., 2017). Moreover, across all three empirical chapters of this dissertation, we found that verbal recall was indeed sensitive to subtle differences in memory for overlapping scene image stimuli. In **Chapter IV**, we went a step further to decompose verbal recall into idiosyncratic and normative components, and found that verbal recall is also sensitive to fine-grained participant-specific effects in memory for scene images. In summary, we showed that complex visual stimuli can be quantified using complex measures such as image segmentation and verbal recall, contributing to a shift in the field to consider how visual memory processes manifest in more naturalistic contexts.

Overlap shapes memory for naturalistic scene images

It is well established that overlapping experiences shape how we form memories by influencing (1) whether those experiences are remembered (Anderson & Spellman, 1995; Antony et al., 2022; Mensink & Raaijmakers, 1988; Osgood, 1949; Staresina et al., 2009), and (2) how those experiences are remembered relative to one another (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024). However, these studies were largely limited to controlled stimuli and simple feature measures (Chanales et al., 2021; Drascher & Kuhl, 2022; Morrow & Clewett, 2025; Tandoc et al., 2024). In this dissertation, we used naturalistic scene images and verbal recall as our measure of memory content. Due to these major differences, it is difficult to make direct comparisons to prior work; instead, we considered whether our results could be theoretically analogous. In **Chapter II** and **Chapter III**, we showed that overlap leads to a separation in verbal recall. These findings likely reflect adaptive mechanisms common across individuals. Meanwhile in **Chapter IV**, we showed that verbal recall of overlapping scenes includes dissociable idiosyncratic and shared components, indicating overlap may be resolved in different ways within and across individuals.

Prior work examining the influence of overlap on memory has demonstrated an adaptive separation, or repulsion, of memories for overlapping stimuli (Chanales et al., 2021; Drascher & Kuhl, 2022). Specifically, these studies used highly controlled stimuli that vary along a specific feature dimension (e.g., color), and showed that memory for that feature becomes biased such that a stimulus is remembered as more different from an overlapping competitor than it actually was (Chanales et al., 2021; Drascher & Kuhl, 2022). A critical limitation of this approach, which reduces generalizability of results, is that the discriminating features of a stimulus are one of the only features along which stimuli could vary. Moreover, memory for each feature must be explicitly probed, automatically orienting participants to those features. To improve generalizability, we used verbal recall as our measure of memory content, wherein participants freely describe images and are naïve to the specific feature dimensions quantified by large language models. Interestingly, in **Chapter II** and **Chapter III** we found separation in verbal recall – analogous to prior accounts of memory repulsion. **Chapter II** demonstrated a separation for multiple overlapping images which, on its own, could simply reflect an increase in item-specific information (i.e., precision; Drascher & Kuhl, 2022). Yet in **Chapter III**, we found a competitor-specific prioritization of diagnostic features in verbal recall. Specifically, dissimilar image patches between an image and its competitor were more important in verbal recall of that image. In summary, we showed that adaptive (repulsion-like) changes exist not only along highly specific stimulus features, but also in verbal recall of scene images.

Chapter II and **Chapter III** considered how overlapping images shape verbal recall by triggering adaptive changes in memory. These adaptive processes are likely shared across individuals – i.e., reflecting common strategies to maximize memory success. However, the specific features in memory that trigger these adaptive changes may not be the same across individuals. For example, prior knowledge and attention to different details may influence how a stimulus is encoded (Kok et al., 2013; Maunsell & Treue, 2006). As such, **Chapter IV** explored how idiosyncratic memory content is shaped by neural representations during encoding. We found that there are indeed idiosyncratic differences in memory encoding – neural representations in visual and parietal cortex differently reflect idiosyncratic and shared features in memory content, respectively. Moreover, hippocampus CA23DG – a region known to adaptively differentiate overlapping memories (Chanales et al., 2017; Dimsdale-Zucker et al., 2018; Favila et al., 2016; Hulbert & Norman, 2015; Kim et al., 2017; Mirjalili et al., 2025; Schlichting et al., 2015; Wanjia

et al., 2021; Wing et al., 2020) – was negatively associated with idiosyncratic memory content. In other words, pairs of stimuli that an individual remembers more similarly will be more differentiated in their own hippocampal activity patterns. As such, idiosyncratic features in memory content may be the very features along which neural representations are anchored to avoid memory interference for overlapping images.

Taken together, we demonstrate idiosyncratic and shared mechanisms by which individuals overcome memory interference for overlapping scene images. In **Chapter II** and **Chapter III**, we demonstrate *shared* mechanisms for *behaviorally* separating memory content of overlapping stimuli. Meanwhile in **Chapter IV**, we demonstrate *neural* differentiation in the hippocampus may anchor to *idiosyncratic* features in memory. These findings are likely highly specific to our stimuli (overlapping scene images) and measure of memory content (verbal recall), so they should be interpreted with caution when generalizing to other contexts. Nevertheless, this work provides insight into the subtle and multidimensional ways in which overlapping experiences influence behavioral and neural expressions of memory.

Conclusion

This dissertation explored how overlapping experiences influence memory outcomes by considering: (1) adaptive differences in verbal recall, and (2) idiosyncratic influences on neural encoding. Using scene image stimuli and verbal recall as our measure of memory content, we considered how overlap influences memory in more naturalistic contexts. We found that overlap leads to a separation in verbal recall across two experiments and two empirical chapters. Moreover, we found that idiosyncratic and shared verbal recall are dissociably represented across cortex, and that idiosyncratic verbal recall may anchor adaptive differentiation of overlapping stimuli in hippocampus. However, the current findings are highly specific to our methods, and would not necessarily generalize to other naturalistic or autobiographical contexts. Rather, this work serves as a starting point, demonstrating that the development of innovative methods to quantify naturalistic memories can yield powerful insight into how overlap shapes complex real-world memories.

References Cited

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, *102*(1), 68–100. https://doi.org/10.1037/0033-295X.102.1.68

Antony, J. W., Romero, A., Vierra, A. H., Luenser, R. S., Hawkins, R. D., & Bennion, K. A. (2022). Semantic relatedness retroactively boosts memory and promotes memory interdependence across episodes. *eLife*, *11*, e72519. https://doi.org/10.7554/eLife.72519

Bae, G.-Y., & Luck, S. J. (2017). Interactions between visual working memory representations. *Attention, Perception & Psychophysics*, *79*(8), 2376–2395. https://doi.org/10.3758/s13414-017-1404-8

Chanales, A. J. H., Oza, A., Favila, S. E., & Kuhl, B. A. (2017). Overlap among Spatial Memories Triggers Repulsion of Hippocampal Representations. *Current Biology*, *27*(15), 2307-2317.e5. https://doi.org/10.1016/j.cub.2017.06.057

Chanales, A. J. H., Tremblay-McGaw, A. G., Drascher, M. L., & Kuhl, B. A. (2021). Adaptive Repulsion of Long-Term Memory Representations Is Triggered by Event Similarity. *Psychological Science*, *32*(5), 705–720. https://doi.org/10.1177/0956797620972490

Chapman, A. F., Chunharas, C., & Störmer, V. S. (2023). Feature-based attention warps the perception of visual features. *Scientific Reports*, *13*(1), 6487. https://doi.org/10.1038/s41598-023-33488-2

Chen, J., Leong, Y. C., Honey, C. J., Yong, C. H., Norman, K. A., & Hasson, U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience*, *20*(1), 115–125. https://doi.org/10.1038/nn.4450

Chunharas, C., Rademaker, R. L., Brady, T. F., & Serences, J. T. (2022). An adaptive perspective on visual working memory distortions. *Journal of Experimental Psychology. General*, *151*(10), 2300–2323. https://doi.org/10.1037/xge0001191

Dimsdale-Zucker, H. R., Ritchey, M., Ekstrom, A. D., Yonelinas, A. P., & Ranganath, C. (2018). CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields. *Nature Communications*, *9*(1), 294. https://doi.org/10.1038/s41467-017-02752-1

Drascher, M. L., & Kuhl, B. A. (2022). Long-term memory interference is resolved via repulsion and precision along diagnostic memory dimensions. *Psychonomic Bulletin & Review*, *29*(5), 1898–1912. https://doi.org/10.3758/s13423-022-02082-4

Favila, S. E., Chanales, A. J. H., & Kuhl, B. A. (2016). Experience-dependent hippocampal pattern differentiation prevents interference during subsequent learning. *Nature Communications*, *7*(1), 11066. https://doi.org/10.1038/ncomms11066

Golomb, J. D. (2015). Divided spatial attention and feature-mixing errors. *Attention, Perception, & Psychophysics*, *77*(8), 2562–2569. https://doi.org/10.3758/s13414-015-0951-0

Hulbert, J. C., & Norman, K. A. (2015). Neural Differentiation Tracks Improved Recall of Competing Memories Following Interleaved Study and Retrieval Practice. *Cerebral Cortex*, *25*(10), 3994–4008. https://doi.org/10.1093/cercor/bhu284

Kim, G., Norman, K. A., & Turk-Browne, N. B. (2017). Neural Differentiation of Incorrectly Predicted Memories. *The Journal of Neuroscience*, *37*(8), 2022–2031. https://doi.org/10.1523/JNEUROSCI.3272-16.2017

Kok, P., Brouwer, G. J., van Gerven, M. A. J., & de Lange, F. P. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *The Journal of Neuroscience*, *33*(41), 16275–16284. https://doi.org/10.1523/JNEUROSCI.0742-13.2013

Magnussen, S., & Greenlee, M. W. (1999). The psychophysics of perceptual memory. *Psychological Research*, *62*(2), 81–92. https://doi.org/10.1007/s004260050043

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. https://doi.org/10.1016/j.tins.2006.04.001

Mensink, G.-J., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review*, *95*(4), 434–455. https://doi.org/10.1037/0033-295X.95.4.434

Mirjalili, S., Wanjia, G., Grätz, D., Wang, E., Mayr, U., & Kuhl, B. A. (2025). Hippocampal transformations occur along dimensions of memory interference. *bioRxiv*, 2025.10.13.682242. https://doi.org/10.1101/2025.10.13.682242

Morrow, E., & Clewett, D. (2025). Distortion of overlapping memories relates to arousal and anxiety. *Cognitive, Affective & Behavioral Neuroscience*, *25*(1), 154–172. https://doi.org/10.3758/s13415-024-01229-1

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review*, *56*(3), 132–143. https://doi.org/10.1037/h0057488

Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(6), 1650–1665. https://doi.org/10.1037/xhp0000110

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications*, *6*(1), 8151. https://doi.org/10.1038/ncomms9151

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event Congruency Enhances Episodic Memory Encoding through Semantic Elaboration and Relational Binding. *Cerebral Cortex (New York, NY)*, *19*(5), 1198–1207. https://doi.org/10.1093/cercor/bhn165

Tandoc, M. C., Dong, C. V., & Schapiro, A. C. (2024). Object Feature Memory Is Distorted by Category Structure. *Open Mind: Discoveries in Cognitive Science*, *8*, 1348–1368. https://doi.org/10.1162/opmi_a_00170

Wang, Y., Lee, H., & Kuhl, B. A. (2023). *Mapping multidimensional content representations to neural and behavioral expressions of episodic memory* [Preprint]. Neuroscience. https://doi.org/10.1101/2023.01.09.523096

Wanjia, G., Favila, S. E., Kim, G., Molitor, R. J., & Kuhl, B. A. (2021). Abrupt hippocampal remapping signals resolution of memory interference. *Nature Communications*, *12*(1), 4816. https://doi.org/10.1038/s41467-021-25126-0

Wing, E. A., Geib, B. R., Wang, W.-C., Monge, Z., Davis, S. W., & Cabeza, R. (2020). Cortical Overlap and Cortical-Hippocampal Interactions Predict Subsequent True and False Memory. *Journal of Neuroscience*, *40*(9), 1920–1930. https://doi.org/10.1523/JNEUROSCI.1766-19.2020