Memory Over Time: Neural Mechanisms for Preserving Memory When Experiences Repeat

by

Futing Zou

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Psychology

Dissertation Committee:

Ben Hutchinson, Chair

Brice Kuhl, Co-Chair

Dasa Zeithamova, Core Member

Luca Mazzucato, Institutional Representative

University of Oregon

Summer 2024

DISSERTATION ABSTRACT

Futing Zou

Doctor of Philosophy in Psychology

Title: Memory Over Time: Neural Mechanisms for Preserving Memory When Experiences Repeat

"Time is what keeps everything from happening at once." ― Ray Cummings

Much of our everyday life involves repeated experiences. The fact that humans are able to form and retain separate memories for experiences that repeat over time is not only important for everyday behavior, but it is foundational to theories of episodic memory. While the effect of repetitions on memory has been the subject of extensive research over decades, the mechanisms that preserve memory across repeated experiences are not well understood.

Across three studies, this dissertation explores how the human brain represents and preserves memories when experiences repeat over time, using measurements of behavior as well as brain activity using functional magnetic resonance imaging (fMRI). In these studies, human participants viewed thousands of visual stimuli that repeated at lags ranging from seconds to many months and then later completed different types of memory probes. In the first chapter, we test how spacing between repeated stimulus encounters with a stimulus influences memory and corresponding neural similarity. We show that the spacing increases the similarity of neural activity patterns in the ventromedial prefrontal cortex and, critically, these increases in neural similarity parallel and predict behavioral benefits of spacing. We demonstrate that these spacing benefits—in brain and behavior—reflect the re-encoding of memories for past experiences. In the second chapter, we characterize the role of neural representations in medial temporal lobe (MTL) subregions in preserving memory for when individual events occurred. We show that the ability to remember the temporal context in which a stimulus originally occurred is predicted by re-

expression of previously-evoked patterns of activity in CA1 and ERC across repeated stimulus encounters, suggesting temporal context reinstatement as a mechanism that protects temporal memories when stimuli repeat. In the final chapter, we develop a novel approach for directly measuring temporal context reinstatement. We show that re-encountering a stimulus reinstates semantic information of temporally-adjacent stimuli that putatively 'compose' the temporal context in which the stimulus was originally encountered. Further, we show that neural similarity in CA1 across repeated stimulus encounters positively relates to the degree of temporal context reinstatement. Collectively, these data highlight the multi-faceted impact of repeated experiences on memory and provide novel insight into the neural mechanisms that preserve memories when experiences repeat across time.

This dissertation includes previously published and unpublished co-authored material.

ACKNOWLEDGMENTS

There are many people that I would like to thank for the help they provided throughout my graduate career. First, I would like to express my deep gratitude to my advisors, Drs. Sarah DuBrow, Ben Hutchinson, and Brice Kuhl, for their invaluable, unwavering support and guidance over the years. It is difficult to overstate how much I have learned from each of them and how grateful I am for the constant support and encouragement they have provided. I would also like to thank other members of my dissertation committee, Drs. Dasa Zeithamova and Luca Mazzucato, for their advice and feedback. Additionally, I would like to thank the many members of the DuBrow, Hutchinson, and Kuhl labs for their tremendous support, friendship, and laughter over the years. Finally, most sincere thanks to my family, especially to my mother Qin Jia, for their endless love and support.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Our everyday experiences often involve people, objects, locations, or actions that repeat over time. For example, you may revisit your favorite coffee shop every weekday, engaging in a routine that includes interacting with the same baristas and ordering your usual drink. Understanding how repetitions of a stimulus influence memory has been the subject of extensive research for more than a century and has fundamentally informed theories of episodic memory. However, a comprehensive understanding of the mechanisms that underlie how memories are represented and maintained across repeated experiences remains elusive. In particular, there are different ways in which repetitions might influence memory.

On the one hand, repetitions of a stimulus are often considered beneficial for memory of that stimulus (Ebbinghaus (1885), 2013; Hintzman, 1976; Karpicke & Roediger, 2008). Indeed, one does not need to research memory to intuit that stimuli repeatedly presented over time are generally remembered better than stimuli presented only once. However, it is not only just the number of repetitions that matters, but the distribution of those repetitions. It is widely acknowledged that spacing out repetitions of a stimulus over time considerably improves memory for that stimulus. This phenomenon, known as the "spacing effect", operates across a wide variety of species (Smolen et al., 2016) and learning material (Cepeda et al., 2006) over long timescales (days, weeks, or months) that has significant implications for memory in everyday and educational settings. While decades of research have produced numerous theories to explain this effect, there is still little agreement on the underlying explanation for why spacing benefits memory. A critical limitation to most leading theories of the spacing effect is that they have not been directly informed

or constrained by evidence of how spacing influences neural representations. In fact, relevant neural evidence is remarkably limited.

On the other hand, however, repetitions of a stimulus also create the potential for interference among memories. For example, the ability to remember the last time you ate pizza may be complicated by memories for the countless other times you have had pizza, or it can be difficult to remember when a particular medication was last taken when it is taken regularly. Each of these examples highlights the significant challenge to a definitional component of episodic memory—temporal memory—which is the ability to remember the time at which individual events occurred (Tulving, 2002). That is, recalling a temporally-specific memory is especially prone to interference because, when an experience is repeated over time, the experience becomes associated with multiple temporal contexts. In other words, the difficulty in remembering the last time you ate pizza is largely attributable to the interference presented by the memories of the many other times you had pizza. The interference among temporal memories is a hallmark of aging and neurocognitive dysfunction (Bellassen et al., 2012; Roberts et al., 2014). While resolving temporal memory interference therefore represents an important component of the episodic memory system, relatively little is known about the neural mechanisms that preserve temporal memory when events are re-encountered.

The aim of this dissertation is to characterize the multi-faceted impact of repeated experiences on memory and ultimately to advance our understanding of how the human brain represents and preserves memories when experiences repeat over time. In particular, the work presented in this dissertation describes a series of experiments that extensively sampled behavioral and functional magnetic resonance imaging (fMRI) expressions of memory to investigate (1) how the distribution of repetitions affects the ability to retain memories over long timescales, (2) how

memory for individual events occurred in time is preserved when events are re-encountered over time, and (3) whether repetitions of a stimulus elicit reinstatement of information from neighboring stimuli of its original encounter. I will review relevant background work in the remaining sections of this chapter before moving on to detailed descriptions of each study in subsequent chapters.

Spaced learning and memory

Cognitive theories of spacing effects

For many types of learning, it is well acknowledged that memory is improved when repetitions of a stimulus are spaced over time, as opposed to massed. Decades of empirical research have produced a variety of different theoretical accounts of spacing effects (Cepeda et al., 2009; Delaney et al., 2010). The most prominent explanations for spacing effects are encoding variability theory and study-phase theory. Encoding variability theory posits that when a stimulus is encountered again after a long delay, it leads to a greater variety of encoding elements being associated with the stimulus (Glenberg, 1979; Melton, 1970; Raaijmakers, 2003). By this account, a more variable representation has more 'points of access' and is therefore more likely to be retrieved (Benjamin & Tullis, 2010). Study-phase retrieval is thought to have an opposing effect: when a stimulus is re-encountered, this triggers retrieval of its original experience and thus strengthen the memory (Hintzman et al., 1975; Thios & D'Agostino, 1976). More recently, many papers argue that spacing effects could arise from some combination of these two processes (Benjamin & Tullis, 2010; Cepeda et al., 2009; Greene, 1989; Landauer, 1969; Raaijmakers, 2003; Siegel & Kahana, 2014). In particular, the two have been proposed in combination to capture the non-monotonic relationship between of spacing and memory—with sufficiently long lags, the benefit of spacing decreases or disappears (Cepeda et al., 2006, 2008, 2009). While these theories

offer plausible explanations for many of the behavioral effects of spacing, no single theoretical account has yet been able to account for all the observed spacing effects because they are not informed, or constrained, by experimental evidence of how spacing influences neural representations. This represents a significant gap in the field, particularly because these leading cognitive theories of spacing effects make specific predictions about underlying representations and the relationship of these representations to subsequent memory.

Neuroimaging studies of spacing effects

Neuroimaging and electrophysiological measures have the potential to uniquely inform theories of spacing effects by comparing neural representations over time. To date, however, evidence is surprisingly limited. A primary issue is that it is difficult, or at least resource intensive, to continuously measure neural representations at the timescales over which spacing effects operate (days, weeks, months). Thus, existing evidence overwhelming comes from studies that only sparsely sample the spacing function. That said, some interesting findings have emerged. For example, at least at short lags (within a single experimental session), there is evidence that greater spacing can actually increase the similarity of neural representations of repeated events (Feng et al., 2019; Glas et al., 2021). In rodents, these increases have specifically been observed in medial prefrontal cortex (Glas et al., 2021). While it is not clear to what extent these increases in neural similarity relate to the behavioral benefits of spacing—particularly over longer timescales—the idea that spacing may benefit memory by increasing similarity is intriguing because it is incongruent with encoding variability theory. It is also currently unclear whether or how spacing-related changes in neural similarity interact with study-phase retrieval.

Repetitions and temporal memory

Temporal context reinstatement across repeated experiences

Whereas the preceding section focuses on how the distribution of repetitions influences whether you will later recognize a stimulus, as noted above, a separate question is how repetitions will influence your ability to remember when, in time, a stimulus occurred.

At the core of temporal context models of episodic memory is the idea that encoding an event involves binding that event to the temporal context in which it was experienced (Howard & Kahana, 2002; Norman et al., 2008). Temporal context representations are often operationalized as a composite of stimuli that neighbor in time as new stimuli are encoded (Howard et al., 2011). Thus, stimuli are encoded in the context of temporally-adjacent stimuli. Putatively, these representations provide a 'time stamp' that can be used to recall, or infer, when a stimulus occurred. Critically, when a past event is retrieved from memory, this is thought to trigger reinstatement of the event's temporal context (Polyn et al., 2009). This idea is strongly motivated by behavioral studies showing that stimuli tend to be recalled from memory alongside stimuli that were encoded nearby in time (Sederberg et al., 2008). Yet, there is a surprising lack of evidence directly linking these putative temporal context representations to temporal memory.

To the extent that stimuli are encoded 'in the context of' other stimuli, how can this be measured? Several fMRI studies have used list-learning paradigms in which an entire list of stimuli is associated with the same context (e.g., a background scene) and have shown that this context is reinstated during subsequent recall (Gershman et al., 2013). However, this approach lacks temporal specificity (an entire list is treated as one temporal context). Taking a different approach, intracranial EEG studies have found that recalling a stimulus from the past reinstates neural activity patterns that preceded and followed the initial encoding of that stimulus (Folkerts et al.,

2018; Manning et al., 2011; Yaffe et al., 2014). However, it is not clear whether these neural activity patterns actually reflect reinstatement of the neighboring stimuli that were encoded in the prior temporal context. Moreover, none of these studies have tested temporal memory.

How might temporal context reinstatement influence temporal memory interference? Consider a stimulus first encountered in one temporal context and later encountered in a new temporal context. This potentially creates a tradeoff such that reinstatement of the original temporal context comes at the expense of encoding the new temporal context (Duncan et al., 2012; Huijbers et al., 2009; Long & Kuhl, 2019). However, in associative learning tasks, human behavioral and neuroimaging studies have found that one way in which associative interference can be mitigated is actually by reinstating older associations during new learning (Chanales et al., 2019; Hulbert & Norman, 2015; Koen & Rugg, 2016; Kuhl et al., 2010; Ritvo et al., 2019). That said, this has not been demonstrated in the domain of temporal memory and therefore raises the interesting possibility that temporal memory interference may likewise be reduced to the extent that prior temporal contexts are reinstated when stimuli re-occur in new temporal contexts.

MTL contributions to temporal context reinstatement and temporal memory

The medial temporal lobe (MTL) has long been implicated in episodic memory, and more recent efforts have focused on identifying how MTL subregions contribute to distinct aspects of memory formation (Squire et al., 2004; Davachi, 2006; Eichenbaum et al., 2007). In particular, within the MTL system, human and rodent studies have implicated hippocampal area CA1 and entorhinal cortex (ERC) as being particularly important for encoding and remembering temporal information (Davachi & DuBrow, 2015; Eichenbaum, 2014; Ranganath & Hsieh, 2016; Tsao et al., 2022). For example, in rodents, so called “time cells” in CA1 have been shown to code for

elapsed time (MacDonald et al., 2011; Mankin et al., 2012; Mau et al., 2018). Likewise, human studies have found that activity patterns in CA1 and ERC represent temporal information (Reddy et al., 2021; Umbach et al., 2020) and activation in these regions has been shown to scale with temporal memory accuracy (Jenkins & Ranganath, 2010; Montchal et al., 2019).

While this surge of work in recent years provide critical support for the roles of these MTL structures, especially the CA1 and ERC, in coding for the temporal information of events, it remains unclear whether or how these representations might preserve temporal memories in the face of repeated stimulus encounters. In particular, there are multiple, mechanistically-distinct ways in which temporal context representations might be formed and expressed in service of temporal memory. On the one hand, when an event is repeated in a new temporal context, encoding the new temporal context as distinct from the prior context may improve discriminability and help protect against interference. On the other hand, motivated by influential models of temporal context and recent neuroimaging studies of associative learning which emphasize the role of reinstatement in memory, when an event is repeated in a new temporal context, this trigger reinstatement of prior experiences, thereby 'preserving' the temporal context information.

Overview of the present work

Researchers have long been interested in understanding how repeated experiences influence memory. The present studies in this dissertation builds on the literature discussed above to answer two central questions: (1) What mechanisms support the benefits of spaced learning? (2) What mechanisms preserve memories of specific temporal contexts across repeated experiences? In particular, this dissertation explores these above questions by using various behavioral methods and pattern-based fMRI analyses to examine neural representations of repeated events across long

timescales and relate these representations to distinct memory expressions. Specifically, I used the Natural Scenes Dataset (Allen et al., 2022), which extensively sampled behavioral and 7T fMRI expressions of memory in eight human participants that each completed 30-40 experimental sessions distributed over the course of almost a year.

In **Chapter II**, I test how spacing between repeated exposures with a stimulus influences memory and corresponding neural similarity. I show that spaced learning increases the similarity of neural activity patterns—specifically within the ventromedial prefrontal cortex. Moreover, I show that these increases in neural similarity parallel and predict the behavioral benefits of spacing. Finally, through a combination of additional behavioral and fMRI analyses, I show that the benefits of spaced learning require a balance between competing processes of memory encoding and memory retrieval, with benefits critically depend on remembering and, in turn, 're-encoding' past experience.

In **Chapter III**, I test whether the similarity (or dissimilarity) of neural representations across repeated stimulus encounters predicts temporal memory. I show that the ability to remember the temporal context in which an event originally occurred is predicted by the re-expression of previously-evoked patterns of activity in CA1 and ERC. Through a rigorous series of analyses, I show that these results strongly align with a temporal context reinstatement account.

In **Chapter IV**, building off of temporal context models and the initial evidence in Chapter III showing that re-expression of activity patterns in CA1 preserves temporal memory, I use natural language processing (NLP) methods and fMRI semantic encoding models to develop a novel approach for directly measuring temporal context reinstatement and test whether temporal context reinstatement relates to CA1 representations. I show that re-encountering a stimulus reinstates semantic information of temporally-adjacent stimuli that putatively 'compose' the temporal

context in which the stimulus was originally encountered. Moreover, I show that neural similarity in CA1 across repeated stimulus encounters positively relates to the degree of temporal context reinstatement. Collectively, these data provide novel insight into the neural mechanisms that preserve memories when experiences repeat across time. **Chapter V** will summarize the presented empirical findings and present avenues for directions for future research.

References Cited

Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., Nau, M., Caron, B., Pestilli, F., Charest, I., Hutchinson, J. B., Naselaris, T., & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. Nature Neuroscience, 25(1), Article 1. https://doi.org/10.1038/s41593-021-00962-x

Bellassen, V., Iglói, K., de Souza, L. C., Dubois, B., & Rondi-Reig, L. (2012). Temporal Order Memory Assessed during Spatiotemporal Navigation As a Behavioral Cognitive Marker for Differential Alzheimer's Disease Diagnosis. The Journal of Neuroscience, 32(6), 1942–1952. https://doi.org/10.1523/JNEUROSCI.4556-11.2012

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? Cognitive Psychology, 61(3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing Distributed Practice. Experimental Psychology, 56(4), 236–246. https://doi.org/10.1027/1618-3169.56.4.236

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

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing Effects in Learning: A Temporal Ridgeline of Optimal Retention. Psychological Science, 19(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

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

Davachi, L. (2006). Item, context and relational episodic encoding in humans. Current Opinion in Neurobiology, 16(6), 693–700. https://doi.org/10.1016/j.conb.2006.10.012

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. Trends in Cognitive Sciences, 19(2), 92–99. https://doi.org/10.1016/j.tics.2014.12.004

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In The psychology of learning and motivation: Advances in research and theory, Vol. 53 (pp. 63–147). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(10)53003-2

Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory's Penumbra: Episodic Memory Decisions Induce Lingering Mnemonic Biases. Science, 337(6093), 485–487. https://doi.org/10.1126/science.1221936

Ebbinghaus (1885), H. (2013). Memory: A Contribution to Experimental Psychology. Annals of Neurosciences, 20(4), 155–156. https://doi.org/10.5214/ans.0972.7531.200408

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. Nature Reviews Neuroscience, 15(11), Article 11. https://doi.org/10.1038/nrn3827

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. Annual Review of Neuroscience, 30(1), 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

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. Journal of Neuroscience, 39(27), 5351–5360. https://doi.org/10.1523/JNEUROSCI.2741-18.2019

Folkerts, S., Rutishauser, U., & Howard, M. W. (2018). Human Episodic Memory Retrieval Is Accompanied by a Neural Contiguity Effect. Journal of Neuroscience, 38(17), 4200–4211. https://doi.org/10.1523/JNEUROSCI.2312-17.2018

Gershman, S. J., Schapiro, A. C., Hupbach, A., & Norman, K. A. (2013). Neural context reinstatement predicts memory misattribution. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 33(20), 8590–8595. https://doi.org/10.1523/JNEUROSCI.0096-13.2013

Glas, A., Hübener, M., Bonhoeffer, T., & Goltstein, P. M. (2021). Spaced training enhances memory and prefrontal ensemble stability in mice. Current Biology, 31(18), 4052-4061.e6. https://doi.org/10.1016/j.cub.2021.06.085

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. Memory & Cognition, 7(2), 95–112. https://doi.org/10.3758/BF03197590

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(3), 371–377. https://doi.org/10.1037/0278-7393.15.3.371

Hintzman, D. L. (1976). Repetition and Memory. In Psychology of Learning and Motivation (Vol. 10, pp. 47–91). Elsevier. https://doi.org/10.1016/S0079-7421(08)60464-8

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). Spacing judgments as an index of study-phase retrieval. Journal of Experimental Psychology: Human Learning and Memory, 1(1), 31–40. https://doi.org/10.1037/0278-7393.1.1.31

Howard, M. W., & Kahana, M. J. (2002). A Distributed Representation of Temporal Context. Journal of Mathematical Psychology, 46(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Howard, M. W., Shankar, K. H., & Jagadisan, U. K. K. (2011). Constructing Semantic Representations From a Gradually Changing Representation of Temporal Context. Topics in Cognitive Science, 3(1), 48–73. https://doi.org/10.1111/j.1756-8765.2010.01112.x

Huijbers, W., Pennartz, C. M., Cabeza, R., & Daselaar, S. M. (2009). When Learning and Remembering Compete: A Functional MRI Study. PLOS Biology, 7(1), e1000011. https://doi.org/10.1371/journal.pbio.1000011

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

Jenkins, L. J., & Ranganath, C. (2010). Prefrontal and Medial Temporal Lobe Activity at Encoding Predicts Temporal Context Memory. The Journal of Neuroscience, 30(46), 15558–15565. https://doi.org/10.1523/JNEUROSCI.1337-10.2010

Karpicke, J. D., & Roediger, H. L. (2008). The Critical Importance of Retrieval for Learning. Science, 319(5865), 966–968. https://doi.org/10.1126/science.1152408

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, 36(15), 4389–4399. https://doi.org/10.1523/JNEUROSCI.4099-15.2016

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

Landauer, T. K. (1969). Reinforcement as consolidation. Psychological Review, 76(1), 82–96. https://doi.org/10.1037/h0026746

Long, N. M., & Kuhl, B. A. (2019). Decoding the tradeoff between encoding and retrieval to predict memory for overlapping events. NeuroImage, 201, 116001. https://doi.org/10.1016/j.neuroimage.2019.07.014

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. Neuron, 71(4), 737–749. https://doi.org/10.1016/j.neuron.2011.07.012

Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., & Leutgeb, J. K. (2012). Neuronal code for extended time in the hippocampus. Proceedings of the National Academy of Sciences, 109(47), 19462–19467. https://doi.org/10.1073/pnas.1214107109

Manning, J. R., Polyn, S. M., Baltuch, G. H., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. Proceedings of the National Academy of Sciences, 108(31), 12893–12897. https://doi.org/10.1073/pnas.1015174108

Mau, W., Sullivan, D. W., Kinsky, N. R., Hasselmo, M. E., Howard, M. W., & Eichenbaum, H. (2018). The Same Hippocampal CA1 Population Simultaneously Codes Temporal Information over Multiple Timescales. Current Biology, 28(10), 1499-1508.e4. https://doi.org/10.1016/j.cub.2018.03.051

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. Journal of Verbal Learning and Verbal Behavior, 9(5), 596–606. https://doi.org/10.1016/S0022-5371(70)80107-4

Montchal, M. E., Reagh, Z. M., & Yassa, M. A. (2019). Precise temporal memories are supported by the lateral entorhinal cortex in humans. Nature Neuroscience, 22(2), 284–288. https://doi.org/10.1038/s41593-018-0303-1

Norman, K. A., Detre, G., & Polyn, S. M. (2008). Computational models of episodic memory. In The Cambridge handbook of computational psychology (pp. 189–225). Cambridge University Press. https://doi.org/10.1017/CBO9780511816772.011

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. Psychological Review, 116(1), 129–156. https://doi.org/10.1037/a0014420

Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. Cognitive Science, 27(3), 431–452. https://doi.org/10.1207/s15516709cog2703_5

Ranganath, C., & Hsieh, L.-T. (2016). The hippocampus: A special place for time. Annals of the New York Academy of Sciences, 1369(1), 93–110. https://doi.org/10.1111/nyas.13043

Reddy, L., Zoefel, B., Possel, J. K., Peters, J., Dijksterhuis, D. E., Poncet, M., Straaten, E. C. W. van, Baayen, J. C., Idema, S., & Self, M. W. (2021). Human Hippocampal Neurons Track Moments in a Sequence of Events. Journal of Neuroscience, 41(31), 6714–6725. https://doi.org/10.1523/JNEUROSCI.3157-20.2021

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

Roberts, J. M., Ly, M., Murray, E., & Yassa, M. A. (2014). Temporal discrimination deficits as a function of lag interference in older adults: Temporal Discrimination and Interference in Aging. Hippocampus, 24(10), 1189–1196. https://doi.org/10.1002/hipo.22303

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. Psychological Review, 115(4), 893–912. https://doi.org/10.1037/a0013396

Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. Journal of Experimental Psychology. Learning, Memory, and Cognition, 40(3), 755–764. https://doi.org/10.1037/a0035585

Smolen, P., Zhang, Y., & Byrne, J. H. (2016). The right time to learn: Mechanisms and optimization of spaced learning. Nature Reviews Neuroscience, 17(2), 77–88. https://doi.org/10.1038/nrn.2015.18

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The Medial Temporal Lobe. Annual Review of Neuroscience, 27(1), 279–306. https://doi.org/10.1146/annurev.neuro.27.070203.144130

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. Journal of Verbal Learning and Verbal Behavior, 15(5), 529–536. https://doi.org/10.1016/0022-5371(76)90047-5

Tsao, A., Yousefzadeh, S. A., Meck, W. H., Moser, M.-B., & Moser, E. I. (2022). The neural bases for timing of durations. Nature Reviews Neuroscience, 23(11), Article 11. https://doi.org/10.1038/s41583-022-00623-3

Tulving, E. (2002). Episodic memory: From mind to brain. Annual Review of Psychology, 53, 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B. E., Sperling, M., & Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. Proceedings of the National Academy of Sciences, 117(45), 28463–28474. https://doi.org/10.1073/pnas.2013250117

Yaffe, R. B., Kerr, M. S. D., Damera, S., Sarma, S. V., Inati, S. K., & Zaghloul, K. A. (2014). Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. Proceedings of the National Academy of Sciences, 111(52), 18727–18732. https://doi.org/10.1073/pnas.1417017112

CHAPTER II

BENEFITS OF SPACED LEARNING ARE PREDICTED BY RE-ENCODING OF PAST EXPERIENCES IN VENTROMEDIAL PREFRONTAL CORTEX

This chapter is submitted for publication by Futing Zou, Brice A. Kuhl, Sarah DuBrow, and J. Benjamin Hutchinson.

Abstract

More than a century of research shows that spaced learning improves long-term memory. Yet, there remains debate concerning why. A major limitation to resolving theoretical debates is the lack of evidence for how neural representations change as a function of spacing. Here, leveraging a massive-scale 7T human fMRI dataset, we tracked neural representations and behavioral expressions of memory as participants viewed thousands of natural scene images that repeated at lags ranging from seconds to many months. We show that spaced learning increases the similarity of human ventromedial prefrontal cortex representations across stimulus encounters and, critically, these increases parallel and predict the behavioral benefits of spacing. Additionally, we show that these spacing benefits critically depend on remembering and, in turn, 're-encoding' past experience. Collectively, our findings provide fundamental insight into how spaced learning influences neural representations and why spacing is beneficial.

Introduction

One of the most robust and well-documented phenomena in human memory research is that when repetitions of a stimulus are spaced over time—as opposed to massed—long-term memory for that stimulus is improved (Ebbinghaus (1885), 2013). This spacing effect operates across a wide variety of memoranda (Cepeda et al., 2006) and species(Smolen et al., 2016) and

also across impressively long timescales (Bahrick et al., 1993; Cepeda et al., 2008, 2009; Küpper-Tetzel et al., 2014). However, despite the ubiquity of this effect and its significant implications for memory in everyday and educational settings, there remains debate concerning why spaced learning improves memory. Why is a repeated stimulus more effectively encoded when more time has elapsed since it was previously encoded?

While a number of theoretical accounts of spacing effects have been advanced (Cepeda et al., 2009; Delaney et al., 2010; Smolen et al., 2016), one of the most prominent ideas is that spaced learning benefits memory by increasing encoding variability (Glenberg, 1979; Melton, 1970; Raaijmakers, 2003). By this account, when a stimulus is re-encountered after a long delay, it is encoded differently than the first time it was encountered. A more variable representation is argued to have more 'points of access' and, therefore, to increase the likelihood of later retrieval (Benjamin & Tullis, 2010). However, if encoding variability were the only factor, memory would monotonically increase with greater spacing (the more spacing, the better). In contrast, studies have shown that as the lag between stimulus repetitions increases, the benefit of spacing will eventually reverse, producing an 'inverted U' shaped relationship between spacing and subsequent memory (Cepeda et al., 2006, 2008, 2009). This non-monotonic pattern is often only observed when long timescales are considered (spacing on the order of weeks or months), but it is extremely informative, at a theoretical level, because it suggests a second factor also contributes to spacing effects—a factor that explains diminished benefits at very long lags. Perhaps the most commonly-advanced second factor is study-phase retrieval (Hintzman et al., 1975; Thios & D'Agostino, 1976). By this account, when a stimulus is re-encountered, it triggers retrieval of the original encounter (benefitting memory), but the benefit of retrieval is negatively related to lag because forgetting of the first encounter becomes more likely with longer lags. While the combination of

encoding variability and study-phase retrieval has high explanatory power (Benjamin & Tullis, 2010; Cepeda et al., 2009; Greene, 1989; Landauer, 1969; Raaijmakers, 2003; Siegel & Kahana, 2014), a fundamental limitation of almost all leading theoretical accounts of spacing effects is that they are not directly informed, or constrained, by experimental evidence of how spacing influences neural representations. This is particularly glaring in the case of encoding variability, which makes obvious predictions about the variability of neural representations over time and the relationship of this variability to subsequent memory.

The lack of integration of neural evidence with cognitive theories of spacing effects is partly due to the surprisingly limited number of neuroimaging / electrophysiological studies that have characterized neural representations as a function of spacing. Moreover, existing evidence is largely limited to studies that measure spacing effects over short timescales (a single experimental session), making it difficult to test ideas from cognitive theories that are specifically motivated by behavioral spacing effects over long timescales. A primary issue is that it is difficult, or at least resource intensive, to measure neural representations across days, weeks, or months. That said, there is recent evidence that, at least at short timescales, neural representations of repeated events may actually be more similar with greater spacing (Feng et al., 2019; Glas et al., 2021)—a finding not predicted or easily explained by an encoding-variability account. However, it is currently unclear why neural representations would become more similar with spacing, how neural similarity might change over longer timescales, and what neural similarity ultimately tells us about why spaced learning is beneficial.

Here, we measured neural representations and behavioral expressions of memory for stimuli repeated over long timescales with the goal of integrating neural and behavioral measures to inform theories of spacing effects. To this end, we leveraged the Natural Scenes Dataset (Allen

et al., 2022)—an unprecedented dataset that combines behavioral measures of memory and 7T fMRI for human participants that studied stimuli distributed across 30-40 experimental sessions over an 8-10 month window (Fig. 2.1a-b). Across these sessions, participants viewed 9,209-10,000 natural scene images presented up to three times each (E1, E2, E3) and made recognition memory decisions at each encounter (continuous recognition task). This yielded a remarkably wide range of lags (spacing) between the first two exposures, from 4 seconds up to 302 days (Fig. 2.1c). Of central interest here was how spacing between the first two exposures with a stimulus (E1-E2 lag) influenced corresponding neural similarity (E1-E2 similarity) and whether neural measures of similarity parallel and predict behavioral benefits of spacing (recognition memory at E3). We focused our analyses on ventromedial prefrontal cortex (vmPFC), motivated by the extensive human neuroimaging literature implicating vmPFC in episodic memory across long timescales (Barry et al., 2018; Bonnici et al., 2012; Gilboa & Moscovitch, 2021; Moscovitch & Nadel, 2019) and by recent evidence of spacing effects on neural similarity in the rodent medial prefrontal cortex (Glas et al., 2021). To preview, we show spacing-related increases in vmPFC similarity that (a) parallel and predict behavioral benefits of spacing, and (b) reflect the re-encoding of memories for prior encounters.

Results

Spaced learning benefits long-term memory

Our primary prediction for behavior was that spacing between the first two exposures (E1-E2 lag) would benefit subsequent memory at the third exposure (E3 recognition). However, prior behavioral work that has considered spacing at long timescales (weeks or months) suggests that memory performance does not monotonically improve as a function of spacing (Cepeda et al.,

2008). Rather, as spacing increases, memory benefits first increase, but eventually decrease (Cepeda et al., 2009; Glenberg, 1976). By some accounts, this non-monotonic function reflects the fact that the benefits of spacing depend on stimuli being recognized when they are re-encountered (recognition at E2 in current study) (Raaijmakers, 2003). From this perspective, the non-monotonic relationship between spacing and memory reflects two competing influences: (i) greater spacing is associated with better memory so long as stimuli are recognized at E2, but (ii) the probability of recognizing a stimulus at E2 decreases with spacing. In the current study, because memory was measured at each encounter (E1, E2, E3), we were able to directly test whether the relationship between E1-E2 spacing (hereafter referred to as spacing) and E3 recognition memory (hereafter referred to as subsequent memory) depended on successful recognition at E2.

An additional nuance to spacing effects is that the optimal amount of spacing (the 'peak' in the non-monotonic function) has been shown to increase as a function of the retention interval (RI; here, the RI is the E2-E3 lag) (Antony et al., 2023; A. S. N. Kim et al., 2019; Küpper-Tetzel et al., 2014). For the sake of comparison with this line of work, we report spacing effects as a function of RI for our initial behavioral analyses. However, because our fMRI analyses focus on how E1-E2 spacing influences E1-E2 neural pattern similarity, the RI is not of direct relevance.

Using mixed-effects logistic regression models (see Methods for details), we first tested for relationships between E1-E2 spacing and subsequent memory (hit versus miss at E3), regardless of whether stimuli were successfully recognized at E2 and regardless of behavioral responses at E1. Note that all spacing analyses used the logarithm of the E1-E2 lag intervals (see Methods). Based on prior studies, we predicted a non-monotonic relationship between spacing and subsequent memory, which we tested for as quadratic trends in the logistic regression models. For

this first set of analyses, we generated six different models corresponding to RIs ranging from <10 minutes (shortest RI) to > 3 months (longest RI).

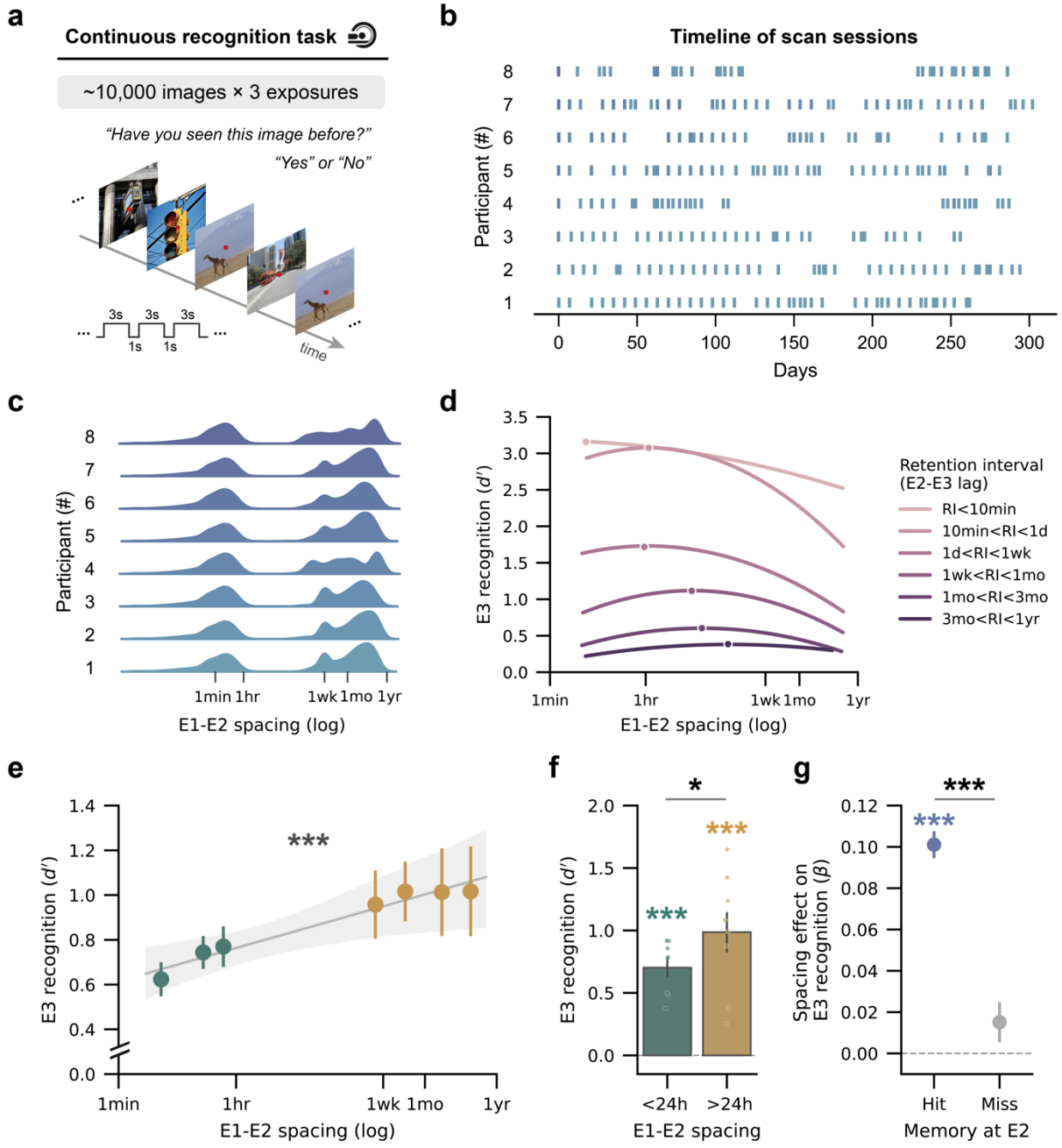

Figure 2.1. Task and behavioral results. a, Trial structure. During fMRI, participants performed a continuous recognition task in which thousands of natural scene images were presented up to three times each with a pseudo-random delay interval between repetitions. **b,** Timeline of 7T fMRI

scan sessions. Each of the eight participants completed 30-40 scan sessions over a 10-month window. The first scan session for each participant corresponds to day 0. **c,** Smoothed histograms showing, for each participant, the relative frequencies of various E1-E2 spacing intervals (from seconds to months), shown on a log scale. **d,** Relationships between spacing (E1-E2 lag) and subsequent memory (E3 recognition) for different retention intervals (RI; E2-E3 lag), using data from all trials (regardless of behavioral responses at E1 and E2). Lines reflect the fits of quadratic trends. Circles mark the peak in the fitted lines. (Also see Fig. 2.5). Qualitatively, the optimal amount of spacing increased as a function of RI. **e,** Subsequent memory (E3 recognition) improved as a function of E1-E2 spacing for images that were successfully recognized at E2 ($\beta = 0.10$, $p < 0.001$, logistic mixed-effects regression). **f,** Subsequent memory performance (E3 recognition) for images with < 24 h spacing and images with > 24 h spacing (conditional on E2 recognition). Memory was above-chance for both groups (ps < 0.001, one-sample t-tests) but was significantly greater for images with > 24 h spacing compared to < 24 h spacing ($t = -2.55$, $p = 0.038$, paired t-test). **g,** The relationship between spacing and subsequent memory (E3 recognition) was significantly stronger for images successfully recognized at E2 (hit) compared to images not recognized at E2 (miss) (hit: $\beta = 0.10$, $p < 0.001$; miss: $p = 0.10$, logistic mixed-effects regression; hit compared to miss: $z = 7.89$, $p < 0.001$, z-test). Throughout the figure, error bars depict mean ± s.e.m.; dots depict independent participants ($n = 8$); *$p < 0.05$; ***$p < 0.001$.

Consistent with prior findings (Küpper-Tetzel et al., 2014; Litman & Davachi, 2008; Rawson & Kintsch, 2005), we found no benefit to spacing for the shortest RI (<10 minutes)—in fact, subsequent memory linearly decreased as a function of spacing, with no evidence of a quadratic trend (see Table 2.1 for all statistics). However, for all RIs greater than 10 minutes (the other five models), we observed significant quadratic trends (ps < 0.001; Table 2.1). Specifically, as spacing increased, subsequent memory first increased, and then decreased. Qualitatively, the 'optimal' amount of spacing (the peak in the quadratic function) increased as a function of the RI (Fig. 2.1d and Fig. 2.5), with the peaks ranging from spacing of ~1 hour to spacing of several days. Most of the models also had significant negative linear trends (Table 2.1).

We next conducted a separate mixed-effects logistic regression model restricted to stimuli correctly recognized at E2 (E2 = 'old' response; hit) and correctly identified as new at E1 (E1 = 'new' response; correct rejection); for this model, RI was treated as a nuisance variable (covariate of no interest). We also excluded stimuli for which the RI was < 24 hours because stimuli that were

successfully recognized at E2 and then tested less than 24 hours later were effectively at ceiling in terms of E3 memory performance (average hit rate across participants = 0.97). This model again yielded a significant quadratic trend (quadratic term: $\beta = -0.01$, $p < 0.001$), but in striking contrast to the negative linear trends observed when ignoring E2 recognition (Table 2.1), there was a robust, positive linear relationship between spacing and subsequent memory ($\beta = 0.10$, $p < 0.001$, logistic mixed-effects regression; Fig. 2.1e). We also directly compared subsequent memory performance (measured as d') for stimuli that had spacing > 24 hours versus spacing < 24 hours, again restricting analysis to stimuli that were correctly recognized at E2. While performance was well above chance for both conditions (< 24 h: $t = 9.21$, $p < 0.001$; > 24 h: $t = 5.80$, $p < 0.001$, one-sample t-tests; Fig. 2.1f), subsequent memory performance was significantly higher for > 24 h spacing compared to < 24 h spacing ($t = -2.55$, $p = 0.038$, paired-samples t-test; Fig. 2.1f).

Finally, we also ran a model that was conditionalized on E2 not being recognized (E2 = miss; E1 = correct rejection). For this model, we did not observe a significant linear relationship between spacing and subsequent memory (linear trend: $p = 0.10$, logistic mixed-effects regression; Fig. 2.1g). Moreover, the linear relationship between spacing and subsequent memory was significantly stronger when E2 was recognized versus not recognized ($z = 7.89$, $p < 0.001$, z-test; Fig. 2.1g). Thus, the benefits of spaced learning were highly dependent on successful recognition at E2.

Spaced learning strengthens stimulus-specific representations in vmPFC

Having found behavioral evidence that spaced learning benefits subsequent memory, we next assessed whether and how E1-E2 spacing influenced representational similarity across encounters (E1-E2 fMRI pattern similarity). We did this using linear mixed-effects models in

which pattern similarity was the dependent measure (see Methods). For these analyses, we did not exclude any trials based on RI. Importantly, however, we did restrict analyses to E1 and E2 trials that were each associated with correct behavioral responses (E1 = correct rejection, E2 = hit) so that any potential relationships between spacing and fMRI pattern similarity were not confounded with behavioral responses. Additionally, to ensure that pattern similarity did not reflect generic cognitive processes, all pattern similarity analyses used a measure of stimulus-specific similarity. That is, for each image we compared 'within-image' pattern similarity (E1 and E2 = same stimulus) to 'across-image' pattern similarity (E1 and E2' = different stimuli; Fig. 2a). Notably, E2' images were selected such that they shared behavioral responses with and were presented in the same sessions as E2 (thus approximately matching for spacing; see Methods for details). Stimulus-specific similarity values greater than zero provide positive evidence for a representation of a specific stimulus.

We focused our analyses on three regions of interest (ROIs): (1) vmPFC, given our a priori prediction that vmPFC representations would be influenced by spacing (Barry et al., 2018; Glas et al., 2021), (2) early visual cortex (EVC) as a control region that would be sensitive to low-level visual information but would not be expected to contribute to or reflect memory-related effects, and (3) motor cortex (M1) as a control region that would not be expected to be sensitive to visual information or memory-related effects.

Given the positive linear relationship between spacing and subsequent memory that we observed in our behavioral analysis of stimuli that were recognized at E2 (Fig. 2.1e), here we tested for similar linear relationships between spacing and (stimulus-specific) pattern similarity. Intuitively, it might be predicted that greater spacing would be associated with lower pattern similarity (more variability). Indeed, EVC exhibited a strong negative relationship between

spacing and pattern similarity ($\beta = -0.002$, $p < 0.001$, linear mixed-effects regression; Fig. 2.2e). In other words, EVC similarity decreased as a function of spacing. In sharp contrast, spacing was positively related to pattern similarity in vmPFC ($\beta = 0.001$, $p = 0.003$, linear mixed-effects regression; Fig. 2.2b). That is, the vmPFC representation at E2 exhibited greater similarity to E1 when the E1-E2 lag was longer. Binning stimuli with < 24 h versus > 24 h spacing confirmed that vmPFC pattern similarity did not differ from zero for < 24 h spacing ($t = 1.71$, $p = 0.13$, one-sample t-test), but was significantly greater than zero for > 24 h spacing ($t = 3.23$, $p = 0.014$, one-sample t-test; Fig. 2.2c). Thus, stimulus-specific representations in vmPFC only emerged when spacing was relatively high (> 24 h). In contrast, EVC exhibited significant pattern similarity for both < 24 h spacing ($t = 8.55$, $p < 0.001$, one-sample t-test) and for > 24 h spacing ($t = 7.99$, $p < 0.001$, one-sample t-test). Thus, EVC did consistently code for stimulus-specific information, even if these representations were weaker with greater spacing. As expected, we did not observe any evidence that pattern similarity in M1 was influenced by spacing ($p = 0.27$, linear mixed-effects regression; Fig. 2.2f), nor did M1 exhibit stimulus-specific representations for stimuli with < 24 h spacing ($t = 0.61$, $p = 0.56$, one-sample t-test) or > 24 h spacing ($t = -0.18$, $p = 0.86$, one-sample t-test).

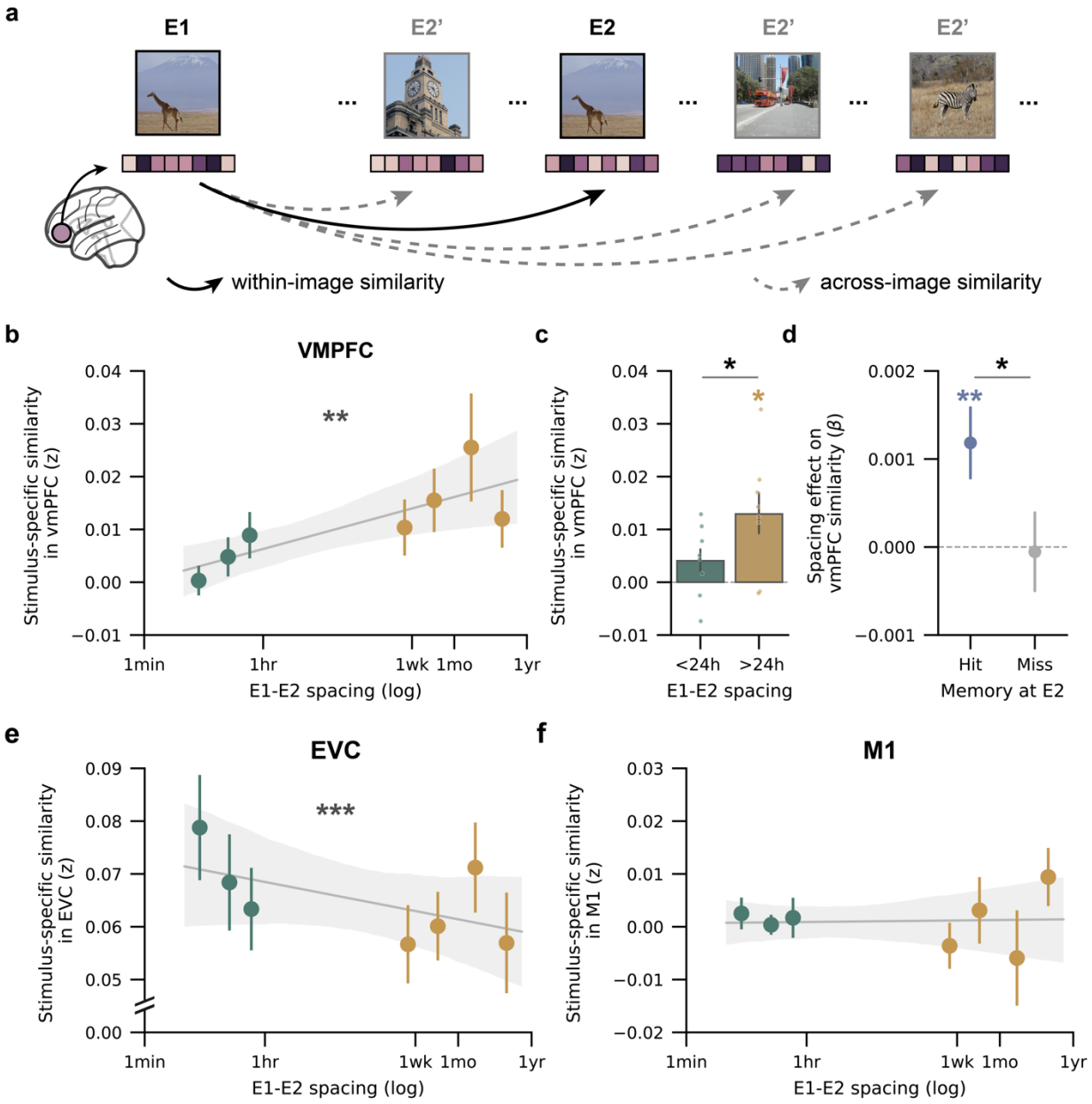

Figure 2.2. Spaced learning strengthens stimulus-specific representations in vmPFC. a, Schematic illustration of stimulus-specific pattern similarity analysis. For each image, we computed pattern similarity scores (z-transformed Pearson correlations) reflecting within-image similarity (E1 and E2 = same stimulus) and across-image similarity (E1 and E2' = different stimuli). Stimuli used to compute the across-image similarity (E2' images) were associated with (1) the same behavioral responses as E2 images and (2) the same session as E2 images (E2 session = E2' session; see Methods for details). For each image, we then computed the difference between within- and across-image similarity, which we refer to as '*stimulus-specific'* pattern similarity. Under this difference measure, values greater than zero indicate positive evidence for a representation of a specific stimulus in a given brain region. **b,** Stimulus-specific similarity in

vmPFC increased as a function of spacing ($\beta = 0.001$, $p = 0.003$, linear mixed-effects regression). **c,** Stimulus-specific similarity values in vmPFC were significantly greater for stimuli with > 24 h spacing compared to < 24 h spacing ($t = -2.82$, $p = 0.026$, paired t-test). Stimulus-specific similarity in vmPFC did not differ from zero for stimuli with < 24 h spacing ($t = 1.71$, $p = 0.13$, one-sample t-test) but was significantly greater than zero for > 24 h spacing ($t = 3.23$, $p = 0.014$, one-sample t-test). Dots depict individual participants ($n = 8$). **d,** The relationship between spacing and vmPFC pattern similarity was stronger (more positive) for images successfully recognized (hit) compared to images not recognized (miss) at E2 (hit: $\beta = 0.001$, $p = 0.003$; miss: $p = 0.90$, linear mixed-effects regression; hit compared to miss: $z = 2.07$, $p = 0.038$, z-test). **e,** Stimulus-specific similarity in EVC decreased as a function of spacing ($\beta = -0.002$, $p < 0.001$, linear mixed-effects regression). **f,** Spacing had no effect on stimulus-specific similarity in M1 ($p = 0.27$, linear mixed-effects regression). Throughout the figure, error bars depict mean ± s.e.m.; *$p < 0.05$; **$p < 0.01$, ***$p < 0.001$.

Importantly, we also confirmed that there was no significant linear relationship between spacing and stimulus-specific similarity in vmPFC when images were not recognized at E2 (p = 0.90, linear mixed-effects regression; Fig. 2.2d). Moreover, the linear relationship between spacing and vmPFC pattern similarity was significantly stronger when stimuli were successfully recognized at E2 compared to when they were not recognized at E2 (z = 2.07, p = 0.038, z-test). Together, these results demonstrate that spaced learning strengthened stimulus-specific representations in vmPFC, but only when stimuli were successfully recognized at E2. These data strongly parallel our behavioral findings: that subsequent memory linearly increased as a function of spacing when stimuli were successfully recognized at E2 (compare Fig. 2.2b to Fig. 2.1e).

It is important to emphasize that conditionalizing the fMRI analyses on successful recognition at E2 avoided potential confounds between spacing and behavioral responses at E2 (namely, as spacing increases there is a lower probability that E2 = hit). Moreover, establishing the relevance of E2 recognition in spacing effects is also of theoretical importance. That said, we also assessed stimulus-specific pattern similarity in vmPFC as a function of spacing for all trials (regardless of behavioral response at E2 or E1). This revealed a significant quadratic trend (quadratic term: β = -0.001, p = 0.04; Fig. 2.6), qualitatively similar to the non-monotonic

relationship between spacing and subsequent memory that we observed in the behavioral analyses that included all trials (Fig. 2.1d). Thus, in multiple respects, we observed a strong parallel between pattern similarity in vmPFC and behavioral effects of spaced learning. Importantly, these parallels were particularly evident—or uniquely observable—because we considered a very wide range of timescales (spacing from seconds to months).

Stimulus-specific similarity in vmPFC predicts behavioral benefits of spacing

We have so far shown that spaced learning (E1-E2 spacing) induced parallel increases in both subsequent memory performance (E3 memory) and stimulus-specific pattern similarity in vmPFC (E1-E2 similarity). We next sought to directly link these behavioral and neural expressions, at the level of individual stimuli. To this end, we used logistic mixed-effects regression models (see Methods), with the dependent variable being subsequent memory and independent variables of pattern similarity and spacing. RI was again included as a covariate and stimuli with RI less than 24 hours were excluded due to ceiling effects in subsequent memory (see Spaced learning benefits long-term memory, above). Primary analyses were restricted to stimuli that were successfully recognized at E2, as in the preceding section.

We first tested for a relationship between pattern similarity in vmPFC and subsequent memory and then tested whether this relationship interacted with spacing. The overall relationship between vmPFC pattern similarity and subsequent memory was not significant ($p = 0.12$, logistic mixed-effects model). However, adding an interaction term to the model revealed that the relationship between vmPFC similarity and subsequent memory strongly depended on spacing ($\beta = 0.14$, $p < 0.001$, logistic mixed-effects regression; Fig. 2.3a). Specifically, the strength of the relationship between vmPFC similarity and subsequent memory increased as a function of spacing.

Follow-up tests confirmed a robust positive relationship between vmPFC pattern similarity and subsequent memory when spacing was > 24 h ($\beta = 0.81$, $p = 0.002$), but not when spacing was < 24 h ($\beta = -0.21$, $p = 0.18$). This complements findings described above (Fig. 2.2b, c) showing that stimulus-specific representations in vmPFC only emerged as spacing increased. Therefore, the increase in stimulus-specific pattern similarity in vmPFC that emerged with greater spacing was clearly linked to subsequent memory.

We next tested whether pattern similarity in the control ROIs—EVC and M1—was related to subsequent memory. Neither region exhibited an overall relationship between pattern similarity and subsequent memory (ps > 0.26), nor did they exhibit an interaction with spacing (ps > 0.64). Thus, despite the presence of strong, stimulus-specific pattern similarity in EVC and a significant change in this pattern similarity with spacing (Fig. 2.2e), EVC representations were not predictive of subsequent memory. To further characterize the selectivity of the relationship, we conducted an exploratory whole-brain analysis (including 360 ROIs based on a cortical atlas (Glasser et al., 2016)) to identify regions in which the relationship between pattern similarity and subsequent memory interacted with spacing. Using an arbitrary threshold of $p = 0.001$, this analysis revealed a single region: left vmPFC (Fig. 2.3d). However, as this exploratory analysis only included cortical areas, we also directly interrogated several medial temporal lobe (MTL) regions that are known to be involved in episodic memory (Davachi et al., 2003; Eichenbaum et al., 2007): hippocampal subfields CA1 and CA2/3/dentate gyrus, entorhinal cortex, perirhinal cortex, and parahippocampal cortex. The relationship between pattern similarity and subsequent memory did not interact with spacing for any of the MTL regions (ps > 0.05).

Finally, we tested whether the observed interaction in vmPFC depended on images being successfully recognized at E2, as with the behavioral spacing effects (Fig. 2.1g) and the spacing-

dependent increase in vmPFC pattern similarity (Fig. 2.2d). Indeed, the interaction in vmPFC was not significant when the regression analysis was restricted to images that were not recognized at E2 (p = 0.55 for interaction term of pattern similarity x spacing on subsequent memory; logistic mixed-effects regression). Thus, at the behavioral and neural levels, the effects of spaced learning depended on successful recognition at E2.

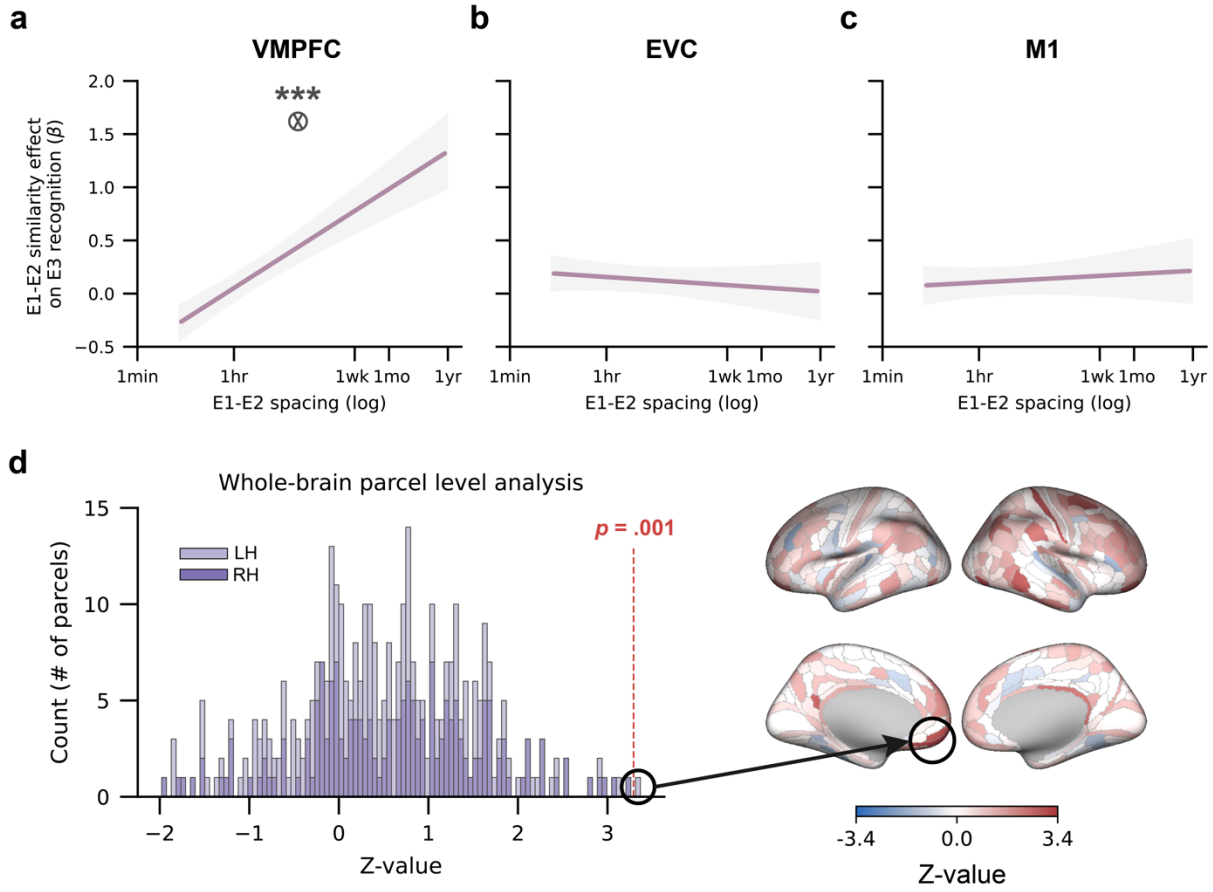

Figure 2.3. Stimulus-specific similarity in vmPFC predicts behavioral benefits of spacing. a, The relationship between vmPFC similarity (E1-E2 similarity) and subsequent memory (E3 recognition) depended on (significantly interacted with) spacing (β = 0.14, $p < 0.001$, logistic mixed-effects regression). **b,** The relationship between EVC similarity and subsequent memory did not depend on spacing ($p = 0.64$, logistic mixed-effects regression). **c,** The relationship between M1 similarity and subsequent memory did not depend on spacing ($p = 0.73$, logistic mixed-effects regression). **d,** Exploratory whole-brain analysis to identify cortical regions in which there was a significant interaction between E1-E2 similarity and spacing in predicting subsequent memory. Z

values refer to the *z*-scores calculated for the interaction term in a logistic mixed-effects regression model. Using an arbitrary threshold of $p < 0.001$, the only region to show a significant interaction was a region in left vmPFC. LH, left hemisphere; RH, right hemisphere. Throughout the figure, error bars depict mean ± s.e.m.; ***$p < 0.001$; ⊗ indicates spacing × pattern similarity interaction.

Stimulus-specific similarity in vmPFC reflects encoding-related processes

Thus far, our findings demonstrate a parallel and direct link between the behavioral benefits of spaced learning and spacing-dependent increases in vmPFC pattern similarity. However, these findings raise the obvious question: why does greater spacing increase pattern similarity in vmPFC? The fact that increases in vmPFC similarity depended on E2 recognition might suggest that vmPFC tracked successful retrieval of the original encounter. However, retrieval strength should not increase with greater spacing—it should decrease. Alternatively, many theoretical accounts of spacing effects emphasize the importance of encoding processes when a stimulus is repeated (Johnston & Uhl, 1976; Wagner et al., 2000). Moreover, to the extent that encoding and retrieval are opposing memory states (Duncan et al., 2012; Hasselmo et al., 2002; Huijbers et al., 2009; Long & Kuhl, 2019; Wheelock & Long, 2024), decreases in retrieval strength may, in fact, directly support stronger encoding processes. Thus, an interesting possibility is that vmPFC similarity reflected the (re)encoding of retrieved E1 representations. To address this idea, we tested whether vmPFC similarity was correlated with encoding-related neural processes (and, for comparison, with retrieval-related processes). Leveraging the extensive neuroimaging literature on episodic memory encoding and retrieval, we identified, from independent meta-analyses, an ROI strongly associated with successful memory encoding and an ROI strongly associated with successful recollection (memory retrieval accompanied by contextual information) (see Methods). We then tested, on a stimulus-by-stimulus basis, whether univariate activation in these ROIs at E2 correlated with the degree of stimulus-specific E1-E2 pattern similarity in vmPFC. This was tested

using linear mixed-effects regression models with vmPFC similarity as the dependent variable and independent variables including univariate activation (either in the encoding or retrieval ROI) and spacing. These analyses were again restricted to stimuli associated with correct recognition at E2.

Strikingly, higher E1-E2 pattern similarity in vmPFC was predicted by greater E2 activation in the encoding-related ROI (left inferior frontal gyrus; $\beta = 0.006$, $p < 0.001$, linear mixed-effects regression; Fig. 2.4a) and by lower E2 activation in the retrieval-related ROI (left angular gyrus; $\beta = -0.003$, $p = 0.040$, linear mixed-effects regression; Fig. 2.4b). Importantly, because these models included spacing as a factor, these results indicate that univariate activation in the encoding and retrieval ROIs explained variance in vmPFC pattern similarity above and beyond that explained by spacing alone.

Exploratory, whole-brain analyses (Fig. 2.7) did not reveal additional ROIs (beyond left inferior frontal gyrus) in which activity positively correlated with vmPFC pattern similarity. However, activity in several additional ROIs (beyond left angular gyrus) was negatively correlated with vmPFC similarity. These regions largely overlapped with the network of brain regions that has been implicated in the recollection of episodic memories (Rugg & Vilberg, 2013; Wagner et al., 2005). Together, these results provide compelling evidence that, for stimuli correctly recognized at E2, greater E1-E2 pattern similarity in vmPFC was related to stronger engagement of encoding-related processes.

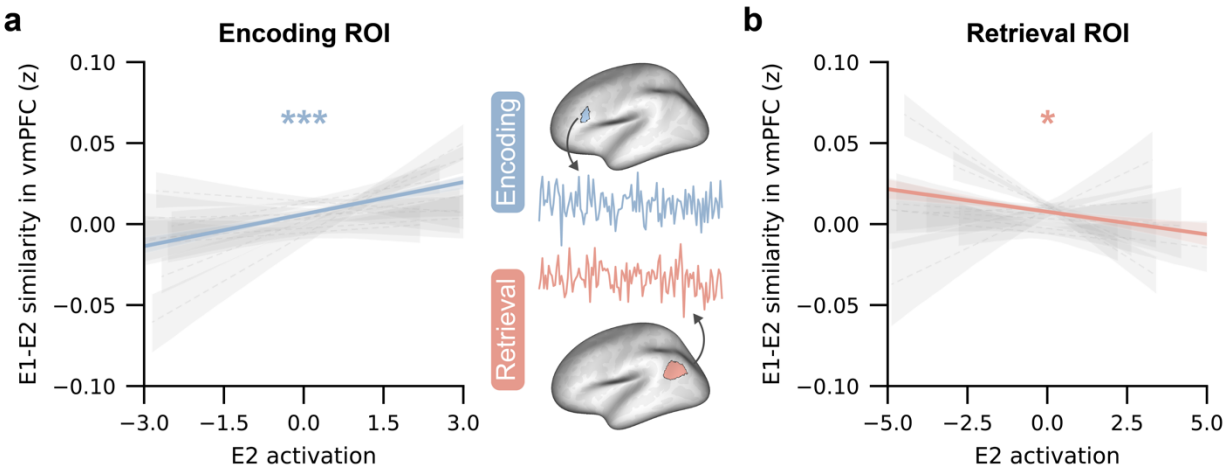

Figure 2.4. Stimulus-specific similarity in vmPFC reflects encoding-related processes. a, There was a positive relationship between stimulus-specific E1-E2 pattern similarity in vmPFC and E2 activation in the encoding ROI (β = 0.007, p < 0.001, linear mixed-effects regression). **b,** There was a negative relationship between pattern similarity in vmPFC and E2 activation in the retrieval ROI (β = -0.003, p = 0.04, linear mixed-effects regression). Throughout the figure, gray dashed lines depict independent participants (n = 8); color lines depict group-level relationship; *p < 0.05; ***p < 0.001.

Discussion

Despite over a century of research related to spacing effects in memory, there remains debate about the underlying explanation for why spacing benefits memory. A major limitation to most current theories of spacing effects is that they have not been directly informed by—or constrained by—evidence from neural measures. Relevant neural evidence is limited, at least in part, because it is difficult to continuously measure neural representations over the timescales across which spacing effects operate (hours, days, weeks, or months). Here, we leveraged a unique human fMRI dataset that allowed us to measure neural representations and behavioral effects of spacing at lags that ranged from seconds to many months. By jointly considering these neural and behavioral expressions—and using each measure to understand the other—we were able to gain fundamental new insight into why spacing benefits memory.

An important finding from behavioral studies of spacing effects is that the relationship between spacing and memory is non-monotonic (Cepeda et al., 2008). Based on this finding, leading theories of spacing effects involve two-factor accounts (Benjamin & Tullis, 2010; Greene, 1989; Maddox, 2016; Raaijmakers, 2003)—one factor to explain the 'rise' in the spacing function (benefits of spacing) and another factor to explain the 'fall' in the spacing function (costs of spacing). Here, we first replicated the classic non-monotonic relationship between spacing and memory when considering data from all trials (without conditionalizing on E2 memory; Fig. 2.1d). The peak in the observed functions (the optimal amount of spacing) fell somewhere between 1 hour and 1 week, depending on the retention interval (Fig. 2.1d and Fig. 2.5). We then observed a qualitatively similar non-monotonic relationship between spacing and vmPFC pattern similarity (E1-E2 similarity; Fig. 2.6), again with a peak that fell between 1 hour and 1 week. At a broad level, these behavioral and fMRI data establish that the current dataset was extremely well suited to studying spacing effects and highlight the value of considering spacing effects over long timescales.

One of the most prominent explanations for the benefits of spaced learning (the 'rise' in the spacing function) is based on the idea of encoding variability (Bjork & Allen, 1970; Cowan et al., 2024; Glenberg, 1976, 1979; Melton, 1970, 1970; Raaijmakers, 2003; Siegel & Kahana, 2014). By this account, greater spacing between stimulus exposures leads to more variable encoding of that stimulus (owing to greater change in the encoding context). In turn, a more variably-encoded stimulus has more points of access and, thereby, is more likely to be remembered. Interestingly, we did find that, in early visual cortex (EVC), stimulus-specific representations became markedly less similar (more variable) as a function of spacing (Fig. 2.2e)—a finding that complements other recent evidence (Deitch et al., 2021; Marks & Goard, 2021; Roth & Merriam, 2023). However,

these EVC effects did not parallel or predict the behavioral benefits of spacing (Fig. 2.3b). In contrast, we found that vmPFC similarity increased (became less variable) with greater spacing and, critically, greater vmPFC similarity positively predicted subsequent memory—particularly at long timescales (Fig. 2.3a). Our finding of spacing-related increases in vmPFC similarity is consistent with recent evidence in humans (Feng et al., 2019) and rodents (Glas et al., 2021) of increases in neural similarity with spacing at short timescales. Our finding that these increases predicted memory aligns with other empirical evidence that neural similarity across stimulus repetitions is positively related to memory (Ward et al., 2013; Xue et al., 2010; Zou et al., 2023) and challenges leading theories of spacing effects that emphasize the role of encoding variability. Thus, we were able to identify opposing effects of spacing on representational similarity (increases and decreases in similarity), but we specifically linked the behavioral benefits of spacing to increases in neural similarity.

Why would vmPFC similarity (or similarity in any brain region) increase as a function of spacing? One factor that was clearly relevant to the increase in vmPFC similarity was whether events were recognized when they were re-encountered (i.e., at E2). Indeed, we observed a strong linear increase in vmPFC similarity when analyses were conditionalized on successful recognition at E2 (Fig. 2.2b-d), but no evidence of a linear increase in similarity when analyses were conditionalized on failed recognition at E2 (Fig. 2.2d). Importantly, this dissociation parallels what we observed in behavior (Fig. 2.1e-g). Thus, our data strongly support the prominent idea that spacing effects depend on study-phase retrieval (Hintzman et al., 1975; Thios & D'Agostino, 1976). That said, while study-phase retrieval was necessary for the benefits of spacing to occur, study-phase retrieval does not, on its own, explain why greater spacing would produce better

memory—in fact, the probability of study-phase retrieval should monotonically decrease with greater spacing.

Motivated by theoretical perspectives arguing that encoding and retrieval are opposing neural states (Hasselmo et al., 2002; O'Reilly & McClelland, 1994), we reasoned that decreases in memory retrieval might be beneficial precisely because they allow for greater memory encoding—or, more specifically, re-encoding of a retrieved representation. We tested this by correlating, in a stimulus-by-stimulus manner, E1-E2 similarity in vmPFC with the degree of E2 univariate activation in regions of interest that have independently been firmly established as being involved in memory encoding and memory retrieval. Indeed, we found that vmPFC similarity was not only positively correlated with encoding activation, but negatively correlated with retrieval activation. Thus, even though vmPFC similarity critically depended on successful retrieval of the original encounter, it was correlated with encoding processes. While this combination may seem contradictory, this is precisely the type of trade-off between opposing influences that is required to explain non-monotonic effects of spacing (in the brain and behavior). Put another way, our findings suggest that memory is likely to benefit when the original experience is successfully retrieved and re-encoded. At very short lags, retrieval may be 'too strong,' thereby preventing successful re-encoding; at very long lags, even though an encoding state may be 'strong,' failure to retrieve the original experience prevents re-encoding. Thus, intermediate lags may be optimal because they allow for a balance between retrieval (of the original exposure) and encoding (of the retrieved memory).

While a re-encoding account provides a parsimonious explanation for our findings—and of spacing effects more generally—a recent computational model suggests a slightly different interpretation that also aligns well with our findings. Namely, Antony et al. argue that variability

triggers the abstraction of similarities across stimulus exposures (Antony et al., 2023). That is, when a stimulus is re-encountered after a relatively long delay, the encoding context is more likely to be different and this difference triggers error-driven learning that strengthens common elements across encounters at the expense of unique elements (i.e., abstraction of similarities). The key point is that this account explains spacing effects in terms of increased neural similarity when events are spaced across time. However, an abstraction account does make a unique prediction about memory for contextual information—including memory for when in time each encounter occurred (temporal memory). Specifically, an abstraction account predicts that, with relatively long spacing, retrieval of the original encounter should actively weaken temporal memory for individual encounters. On the one hand, this prediction is challenged by existing behavioral (Hintzman et al., 1975) and fMRI evidence (Zou et al., 2023) indicating that retrieval of past encounters actually enhances temporal memory. On the other hand, it is notable that temporal memory does not necessarily benefit from spaced learning (Zou et al., 2023). Thus, in future work it will be of interest to further characterize the relationship between spaced learning and contextual memory (including temporal memory) as understanding this relationship will help refine theoretical accounts.

The fact that our findings specifically implicate vmPFC in representing event similarity over time is notable in light of the extensive literature demonstrating vmPFC involvement in the integration of information across encoding events (van Kesteren et al., 2012; Zeithamova et al., 2012) and in the formation of schemas (Gilboa & Marlatte, 2017; Tse et al., 2011). In particular, vmPFC is thought to support the encoding of new events into existing schematic representations (Gilboa & Marlatte, 2017; van Kesteren et al., 2012)—a function which resembles the re-encoding account proposed above. Other studies have also implicated vmPFC in the retrieval of remote

memories (Barry et al., 2018; Bonnici et al., 2012; Ezzyat et al., 2018; Frankland & Bontempi, 2005; Gilboa & Moscovitch, 2021; Moscovitch & Nadel, 2019) and in memory consolidation (Gilboa & Moscovitch, 2021; Nieuwenhuis & Takashima, 2011; Takashima et al., 2006). Notably, some of this evidence specifically relates to vmPFC activation increasing as a function of the age of a retrieved memories (Nieuwenhuis & Takashima, 2011; Takashima et al., 2006)—with the idea being that memories 'move' to (or emerge within) vmPFC over time. In contrast, we show that when a stimulus is re-encountered after a relatively long lag, this can actually increase the similarity to the original experience. Thus, our findings are not readily consistent with an account where vmPFC representations only emerge over long timescales. That said, other theories of consolidation allow for the possibility that vmPFC representations are formed immediately, but the reliance on these representations changes over time (Gilboa & Moscovitch, 2021).

In summary, by considering learning events that were spaced from seconds to many months apart, we show that the behavioral benefits of spaced learning are strongly paralleled—and predicted by—the similarity with which vmPFC represented stimuli across exposures. Moreover, through a number of complementary analyses, we show that the benefits of spaced learning—in brain and behavior—reflect a balance between retrieval and encoding processes. Namely, whereas retrieval of an original encounter is a necessary condition for the benefits of spaced learning, it is also important that retrieved information is re-encoded. From this perspective, non-monotonic relationships between spacing and memory can be explained by trade-offs between these two computations.

Methods

Overview

This study reports findings using the Natural Scenes Dataset (NSD; http://naturalscenesdataset.org). The NSD is a large-scale fMRI dataset in which participants performed a continuous recognition task on thousands of color natural scenes over the course of 30–40 high-resolution fMRI (7T) scan sessions. The results in this study are based on data from all of these sessions, from all participants who took part in the NSD study. For a detailed description of the dataset, including methods involved in preprocessing the fMRI data, please see the original data publication(Allen et al., 2022). Below we outline the specific methods relevant to the current study.

Natural Scenes Dataset

Participants

Eight participants from the University of Minnesota community participated in the NSD study (two self-identified males and six self-identified females; age range 19–32 years). All participants were right-handed with no known cognitive deficits nor color blindness and with normal or corrected-to-normal vision. Participants were naïve to the design of the NSD experiment and were not involved in the design or planning of the study. Informed written consent was obtained from all participants before the start of the study, and the experimental protocol was approved by the University of Minnesota Institutional Review Board.

Design and procedure

A detailed description of the experimental design has been reported in the original data publication (Allen et al., 2022). Briefly, participants performed a continuous recognition task in

which they reported whether the current image had been seen at any previous point in the experiment ('old') or if they had not encountered it before ('new'). For each participant, the experiment was split across 40 scan sessions in which 10,000 distinct color natural scenes would be presented three times spaced pseudo-randomly over the course of the entire experiment. Each scanning session consisted of 12 runs (750 trials per session). Each trial lasted 4 s and consisted of the presentation of an image for 3 s and a following 1-s gap (Fig. 2.1a). Participants were able to respond during the entire 4-s period and were also permitted to make multiple responses per trial in cases where they changed their mind. As trials with multiple responses potentially captured more complex cognitive operations related to decision making, we opted to exclude those trials from all analyses.

Participants completed up to 40 scan sessions each and were scanned approximately once a week over the course of 10 months (Fig. 2.1b; the first scan session for each participant corresponds to Day 0). Four of the participants completed the full set of 40 NSD scan sessions. Due to constraints on participant and scanner availability, two participants completed 30 sessions and two participants completed 32 sessions. Accordingly, each participant viewed a total of 9,209-10,000 unique images across 22,500-30,000 trials.

Stimuli

Images used in NSD were taken from the Microsoft Common Objects in Context (COCO) database (Lin et al., 2014). A total of 73,000 unique images were prepared with the intention that each participant would be exposed to 10,000 distinct images (9,000 unique images and 1,000 shared images across participants) three times each across the 40 scan sessions. Image exposures were pseudo-randomly distributed across sessions over the course of almost a year. The presentation structure was determined in advance and fixed across participants so that difficulty of

the recognition task was roughly similar across participants. Distribution of image presentations was controlled to ensure that both short-term and long-term re-exposures were probed. To provide a sense of the overall experimental design, the mean number of distinct images shown once, twice, and all three times within a typical session is 437, 106, and 34, respectively.

As the current study focused on how the spacing between the first two exposures to an image influenced recognition at the third exposure, we only considered images that were presented all three times across the entire experiment. Here, we labeled each exposure based on its presentation order (i.e., the first exposure is E1, the second exposure is E2, and the third exposure is E3). Of critical interest, the spacing between the first two exposures (E1-E2 lag) ranged from 4 seconds to 302 days (see Fig. 2.1c for distribution of the E1-E2 spacing for each participant).

fMRI data acquisition and preprocessing

MRI data was collected at the Center for Magnetic Resonance Research at the University of Minnesota. Imaging was performed on a 7T Siemens Magnetom passively-shielded scanner with a single-channel-transmit, 32-channel-receive RF head coil. Functional images were acquired using whole-brain gradient-echo echo-planar imaging (EPI) at 1.8-mm resolution and 1.6-s repetition time.

Details of the preprocessing of anatomical and functional data are provided in the original data publication (Allen et al., 2022). Briefly, functional data were preprocessed by performing one temporal resampling to correct for slice time differences and one spatial resampling to correct for head motion within and across scan sessions, EPI distortion, and gradient non-linearities. Informed by the original publication, the current study used the 1.0-mm volume preparation of the functional time-series data and "version 2" of the NSD single-trial betas.

Data Analysis

Based on the observation that stimuli that were successfully recognized at E2 and then tested less than 24 hours later (E2-E3 lag < 24 hours) were effectively at ceiling in terms of E3 memory performance (average hit rate across participants: 0.97), all analyses in the current study related to E3 memory only excluded stimuli for which the E2-E3 lag was > 24 hours.

Behavioral data analyses

We first separately tested for relationships between spacing (E1-E2 lag) and subsequent memory (E3 memory) for different retention intervals (RI; E2-E3 lag), regardless of whether stimuli were successfully recognized at E2 (and regardless of behavioral responses at E1). To do so, we generated six different models corresponding to RIs ranging from <10 minutes (shortest retention interval) to > 3 months (longest retention interval). For each RI, we performed the mixed-effects logistic regression analyses that predicted subsequent memory from spacing using both linear and quadratic fits. Specifically, we used a mixed-effects logistic regression model that predicted subsequent memory (hit = 'old' response, miss = 'new' response) from spacing while including (controlling for) several additional factors. These added factors of no interest included the lag between the beginning of the first trial in the experiment and the first exposure (i.e., E1 onset), the retention interval (i.e., E2-E3 lag), and false alarm rates of sessions in which each exposure occurred. All models were constructed with random intercepts for each participant. Because memory is observed to abide by an exponential rule rather than linear time (Wixted, 2004), all temporal lag information (i.e., E1 onset, E1-E2 lag, and E2-E3 lag) was quantified by expressing time intervals in seconds and transforming these intervals with the natural logarithm.

To directly test whether the relationship between spacing and subsequent memory (and fMRI pattern similarity) depended on successful recognition at E2, we first ran analyses restricted to stimuli correctly recognized at E2 (E2 = hit) and correctly rejected at E1 (E1 = 'new' response; correct rejection) and then ran analyses conditionalized on E2 not being successful recognition at E2 (E2 = miss; E1 = correct rejection).

Regions of interest (ROIs) definition

To probe whether the spacing showed a modulation of stimulus-specific representations, a region-of-interest (ROI) analysis was performed. Motivated by prior evidence implicating vmPFC in episodic memory across long timescales (Barry et al., 2018; Bonnici et al., 2012; Gilboa & Moscovitch, 2021; Moscovitch & Nadel, 2019) and by recent evidence of spacing effects on neural similarity in the rodent medial prefrontal cortex (Glas et al., 2021), this analysis was focused on the bilateral ventromedial prefrontal cortex (vmPFC), along with two control ROIs: bilateral early visual cortex (EVC) and bilateral motor cortex (M1). All cortical ROIs were drawn from the surface-based Human Connectome Project multimodal parcellation (HCP-MMP) atlas of human cortical areas (Glasser et al., 2016). An additional exploratory analysis was conducted throughout the whole brain using all parcels available in the HCP-MMP atlas. Based on the established role of the medial temporal lobe (MTL) in human memory, we also repeated certain analyses in a set of subregions there. These MTL ROIs included bilateral CA1, CA2/3/dentate gyrus, entorhinal cortex, perirhinal cortex, and parahippocampal cortex, all manually drawn on the high-resolution T2 images obtained for each participant.

To probe specific hypotheses concerning encoding/retrieval-related processes, a priori ROIs were chosen based on past work. Specifically, an ROI representative of encoding-related

effects was identified using the global peak coordinates supporting encoding success from an independent meta-analysis on the subsequent memory effect (H. Kim, 2011). We then projected the coordinates to the HCP-MMP atlas and used the delimiting parcel as the encoding ROI. Similarly, for an ROI representative of retrieval-related effects, we used the peak coordinates supporting recollection success from another independent meta-analysis of episodic memory retrieval (Vilberg & Rugg, 2008) and identified the delimiting parcel in the HCP-MMP atlas as the retrieval ROI.

Pattern similarity analyses

Pattern similarity was calculated as the Pearson correlation between activity patterns evoked during different image exposures for each ROI. Correlations were z-transformed (Fisher's z) before further analyses were performed. To avoid potential contamination from BOLD signal autocorrelation, all pattern similarity analyses were performed by correlating activity patterns for stimuli across run (i.e., correlations were never performed within the same scanning run).

For our primary analyses related to pattern similarity between E1 and E2, of critical interest was the stimulus-specific pattern similarity. Specifically, for each image ('target'), we compared 'within-image' pattern similarity (E1 and E2 = same stimulus) to 'across-image' pattern similarity (E1 and E2' = different stimuli; Fig. 2.2a). The E2' images used to compute across-image similarity were chosen from the same set of sessions (but different scanning runs) as the target image's E1 and E2, thus controlling for differences in spacing. Further, E2' images were chosen such that they had the same memory outcomes at the first two exposures as the target image. For target images with multiple possible E2' images based on the criteria above, the median value of the across-

image similarity scores was used. The across-image similarity was then subtracted from within-image similarity to yield a stimulus-specific similarity measure of E1-E2 similarity for each image.

Unless indicated otherwise, our main fMRI analyses were restricted to stimuli that were associated with correct behavioral responses at both E1 and E2 (i.e., E1 = 'new' responses, E2 = 'old' responses) so that any potential relationships between spacing and neural pattern similarity were not confounded with behavioral responses.

Statistical analyses

Behavioral and fMRI data were analyzed using a combination of paired t tests and mixed-effects regression models. Relationships between spacing, stimulus-specific pattern similarity and subsequent memory were tested with mixed-effects regression models. For all mixed-effects regression models, we used the participant as a random effect and other variables as fixed effects. All t-tests were two-tailed. A threshold of $p < 0.05$ was used to establish statistical significance for all analyses unless otherwise specified. fMRI analyses were corrected for multiple comparisons when applicable.

Table 2.1. Statistics for relationships between E1-E2 spacing and subsequent E3 memory for different retention intervals (E2-E3 lag) using both linear and quadratic trend analyses for all trials (i.e., regardless of behavioral responses at E1 and E2). To compare the fits of models for each RI, we used ANOVA to test whether adding the quadratic term of spacing led to significantly better prediction of subsequent memory relative to models without the quadratic term.

Retention interval (RI)	Linear fits	Quadratic fits (quadratic term)	Model comparison
RI<10min	β = -0.130 $p < 0.001$***	β = -0.004 $p = 0.425$	$\chi^2 = 0.623$ $p = 0.430$
10min<RI<1d	β = -0.157 $p < 0.001$***	β = -0.017 $p < 0.001$***	$\chi^2 = 25.146$ $p < 0.001$***

1d<RI<1wk	β = -0.113 p < 0.001***	β = -0.023 p < 0.001***	χ^2 = 53.789 p < 0.001***
1wk<RI<1mo	β = -0.045 p < 0.001***	β = -0.022 p < 0.001***	χ^2 = 70.151 p < 0.001***
1mo<RI<3mo	β = -0.012 p = 0.079	β = -0.013 p < 0.001***	χ^2 = 25.021 p < 0.001***
3mo<R1<1yr	β = 0.016 p = 0.007**	β = -0.009 p < 0.001***	χ^2 = 13.965 p < 0.001***

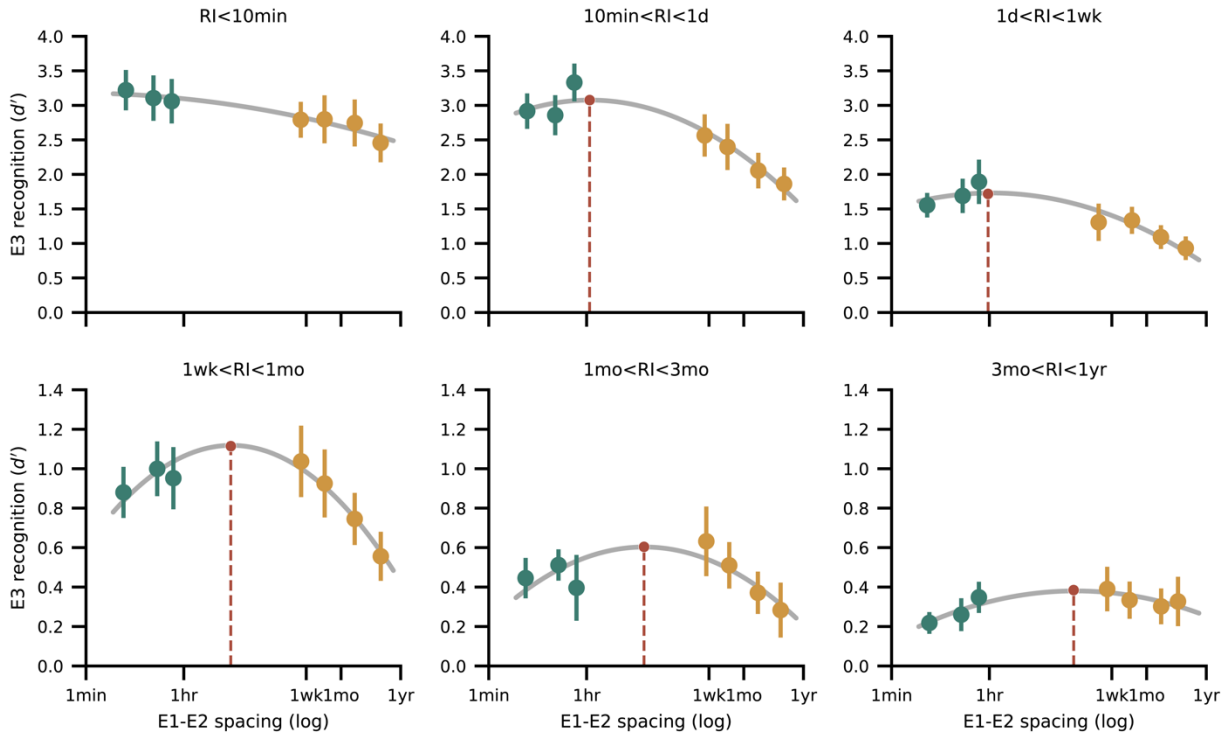

Figure 2.5. Relationships between spacing and subsequent memory (recognition accuracy) for different retention intervals (RIs) for all trials (regardless of whether stimuli were successfully recognized at E2, and regardless of responses at E1).

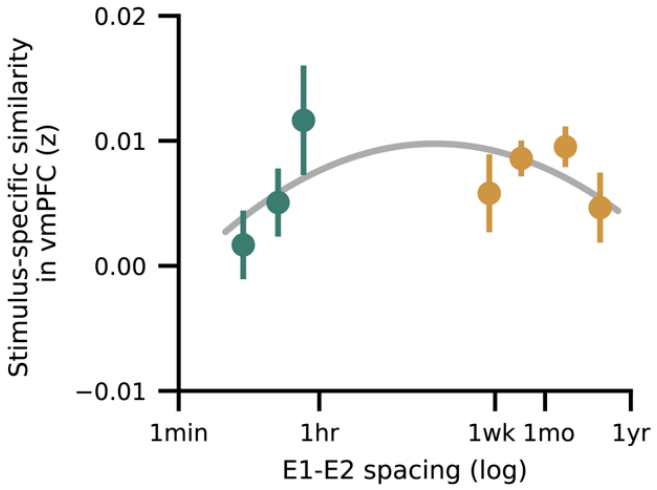

Figure 2.6. Relationship between spacing and stimulus-specific pattern similarity in vmPFC for all trials (regardless of whether stimuli were successfully recognized at E2, and regardless of responses at E1).

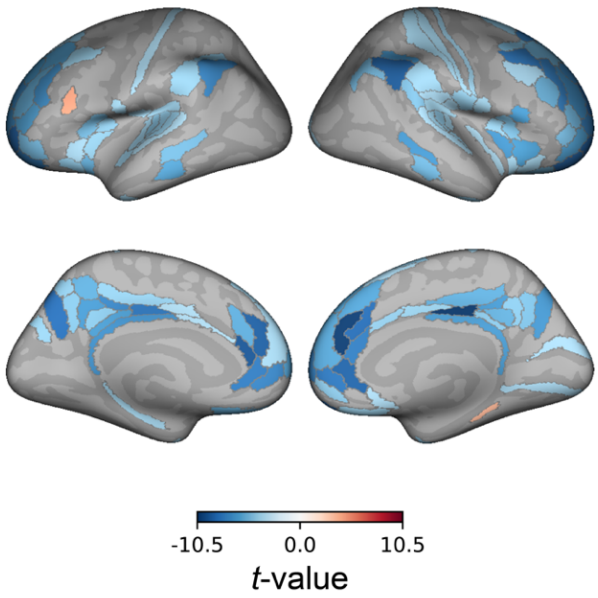

Figure 2.7. Exploratory whole-brain analysis. Linear mixed-effects regression analyses were performed across the whole brain to test for relationships between E2 univariate activation in cortical regions and E1-E2 pattern similarity in vmPFC. Warm colors reflect regions in which E2 univariate activation positively correlated with E1-E2 similarity in vmPFC. Cool colors reflect regions in which E2 univariate activation negatively correlated with E1-E2 similarity in vmPFC.

References Cited

Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., Nau, M., Caron, B., Pestilli, F., Charest, I., Hutchinson, J. B., Naselaris, T., & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nature Neuroscience*, *25*(1), Article 1. https://doi.org/10.1038/s41593-021-00962-x

Antony, J., Liu, X. L., Zheng, Y., Ranganath, C., & O'Reilly, R. C. (2023). *Memory out of context: Spacing effects and decontextualization in a computational model of the medial temporal lobe* (p. 2022.12.01.518703). bioRxiv. https://doi.org/10.1101/2022.12.01.518703

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. E. (1993). Maintenance of Foreign Language Vocabulary and the Spacing Effect. *Psychological Science*, *4*(5), 316–321. https://doi.org/10.1111/j.1467-9280.1993.tb00571.x

Barry, D. N., Chadwick, M. J., & Maguire, E. A. (2018). Nonmonotonic recruitment of ventromedial prefrontal cortex during remote memory recall. *PLOS Biology*, *16*(7), e2005479. https://doi.org/10.1371/journal.pbio.2005479

Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology*, *61*(3), 228–247. https://doi.org/10.1016/j.cogpsych.2010.05.004

Bjork, R. A., & Allen, T. W. (1970). The spacing effect: Consolidation or differential encoding? *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 567–572. https://doi.org/10.1016/S0022-5371(70)80103-7

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

Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing Distributed Practice. *Experimental Psychology*, *56*(4), 236–246. https://doi.org/10.1027/1618-3169.56.4.236

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

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing Effects in Learning: A Temporal Ridgeline of Optimal Retention. *Psychological Science*, *19*(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Cowan, E. T., Zhang, Y., Rottman, B. M., & Murty, V. P. (2024). The effects of mnemonic variability and spacing on memory over multiple timescales. *Proceedings of the National Academy of Sciences*, *121*(12), e2311077121. https://doi.org/10.1073/pnas.2311077121

Davachi, L., Mitchell, J. P., & Wagner, A. D. (2003). Multiple routes to memory: Distinct medial temporal lobe processes build item and source memories. *Proceedings of the National Academy of Sciences*, *100*(4), 2157–2162. https://doi.org/10.1073/pnas.0337195100

Deitch, D., Rubin, A., & Ziv, Y. (2021). Representational drift in the mouse visual cortex. *Current Biology*, *31*(19), 4327-4339.e6. https://doi.org/10.1016/j.cub.2021.07.062

Delaney, P. F., Verkoeijen, P. P. J. L., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. In *The psychology of learning and motivation: Advances in research and theory, Vol. 53* (pp. 63–147). Elsevier Academic Press. https://doi.org/10.1016/S0079-7421(10)53003-2

Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory's Penumbra: Episodic Memory Decisions Induce Lingering Mnemonic Biases. *Science*, *337*(6093), 485–487. https://doi.org/10.1126/science.1221936

Ebbinghaus (1885), H. (2013). Memory: A Contribution to Experimental Psychology. *Annals of Neurosciences*, *20*(4), 155–156. https://doi.org/10.5214/ans.0972.7531.200408

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annual Review of Neuroscience*, *30*(1), 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Ezzyat, Y., Inhoff, M. C., & Davachi, L. (2018). Differentiation of Human Medial Prefrontal Cortex Activity Underlies Long-Term Resistance to Forgetting in Memory. *Journal of Neuroscience*, *38*(48), 10244–10254. https://doi.org/10.1523/JNEUROSCI.2290-17.2018

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. *Journal of Neuroscience*, *39*(27), 5351–5360. https://doi.org/10.1523/JNEUROSCI.2741-18.2019

Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Reviews Neuroscience*, *6*(2), 119–130. https://doi.org/10.1038/nrn1607

Gilboa, A., & Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. *Trends in Cognitive Sciences*, *21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

Gilboa, A., & Moscovitch, M. (2021). No consolidation without representation: Correspondence between neural and psychological representations in recent and remote memory. *Neuron*, *109*(14), 2239–2255. https://doi.org/10.1016/j.neuron.2021.04.025

Glas, A., Hübener, M., Bonhoeffer, T., & Goltstein, P. M. (2021). Spaced training enhances memory and prefrontal ensemble stability in mice. *Current Biology*, *31*(18), 4052-4061.e6. https://doi.org/10.1016/j.cub.2021.06.085

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), Article 7615. https://doi.org/10.1038/nature18933

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 1–16. https://doi.org/10.1016/S0022-5371(76)90002-5

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition*, *7*(2), 95–112. https://doi.org/10.3758/BF03197590

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(3), 371–377. https://doi.org/10.1037/0278-7393.15.3.371

Hasselmo, M. E., Bodelón, C., & Wyble, B. P. (2002). A Proposed Function for Hippocampal Theta Rhythm: Separate Phases of Encoding and Retrieval Enhance Reversal of Prior Learning. *Neural Computation*, *14*(4), 793–817. https://doi.org/10.1162/089976602317318965

Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). Spacing judgments as an index of study-phase retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, *1*(1), 31–40. https://doi.org/10.1037/0278-7393.1.1.31

Huijbers, W., Pennartz, C. M., Cabeza, R., & Daselaar, S. M. (2009). When Learning and Remembering Compete: A Functional MRI Study. *PLOS Biology*, *7*(1), e1000011. https://doi.org/10.1371/journal.pbio.1000011

Johnston, W. A., & Uhl, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(2), 153–160. https://doi.org/10.1037/0278-7393.2.2.153

Kim, A. S. N., Wong-Kee-You, A. M. B., Wiseheart, M., & Rosenbaum, R. S. (2019). The spacing effect stands up to big data. *Behavior Research Methods*, *51*(4), 1485–1497. https://doi.org/10.3758/s13428-018-1184-7

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

Küpper-Tetzel, C. E., Kapler, I. V., & Wiseheart, M. (2014). Contracting, equal, and expanding learning schedules: The optimal distribution of learning sessions depends on retention interval. *Memory & Cognition*, *42*(5), 729–741. https://doi.org/10.3758/s13421-014-0394-1

Landauer, T. K. (1969). Reinforcement as consolidation. *Psychological Review*, *76*(1), 82–96. https://doi.org/10.1037/h0026746

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740–755). Springer.

Litman, L., & Davachi, L. (2008). Distributed learning enhances relational memory consolidation. *Learning & Memory*, *15*(9), 711–716. https://doi.org/10.1101/lm.1132008

Long, N. M., & Kuhl, B. A. (2019). Decoding the tradeoff between encoding and retrieval to predict memory for overlapping events. *NeuroImage*, *201*, 116001. https://doi.org/10.1016/j.neuroimage.2019.07.014

Maddox, G. B. (2016). Understanding the underlying mechanism of the spacing effect in verbal learning: A case for encoding variability and study-phase retrieval. *Journal of Cognitive Psychology*, *28*(6), 684–706. https://doi.org/10.1080/20445911.2016.1181637

Marks, T. D., & Goard, M. J. (2021). Stimulus-dependent representational drift in primary visual cortex. *Nature Communications*, *12*(1), 5169. https://doi.org/10.1038/s41467-021-25436-3

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, *9*(5), 596–606. https://doi.org/10.1016/S0022-5371(70)80107-4

Moscovitch, M., & Nadel, L. (2019). Sculpting Remote Memory: Enduring Hippocampal Traces and vmPFC Reconstructive Processes. *Trends in Cognitive Sciences*, *23*(8), 634–635. https://doi.org/10.1016/j.tics.2019.05.001

Nieuwenhuis, I. L. C., & Takashima, A. (2011). The role of the ventromedial prefrontal cortex in memory consolidation. *Behavioural Brain Research*, *218*(2), 325–334. https://doi.org/10.1016/j.bbr.2010.12.009

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

Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science*, *27*(3), 431–452. https://doi.org/10.1207/s15516709cog2703_5

Rawson, K. A., & Kintsch, W. (2005). Rereading Effects Depend on Time of Test. *Journal of Educational Psychology*, *97*(1), 70–80.

Roth, Z. N., & Merriam, E. P. (2023). Representations in human primary visual cortex drift over time. *Nature Communications*, *14*(1), 4422. https://doi.org/10.1038/s41467-023-40144-w

Rugg, M. D., & Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Current Opinion in Neurobiology*, *23*(2), 255–260. https://doi.org/10.1016/j.conb.2012.11.005

Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(3), 755–764. https://doi.org/10.1037/a0035585

Smolen, P., Zhang, Y., & Byrne, J. H. (2016). The right time to learn: Mechanisms and optimization of spaced learning. *Nature Reviews Neuroscience*, *17*(2), 77–88. https://doi.org/10.1038/nrn.2015.18

Takashima, A., Petersson, K. M., Rutters, F., Tendolkar, I., Jensen, O., Zwarts, M. J., McNaughton, B. L., & Fernández, G. (2006). Declarative memory consolidation in humans: A prospective functional magnetic resonance imaging study. *Proceedings of the National Academy of Sciences*, *103*(3), 756–761. https://doi.org/10.1073/pnas.0507774103

Thios, S. J., & D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, *15*(5), 529–536. https://doi.org/10.1016/0022-5371(76)90047-5

Tse, D., Takeuchi, T., Kakeyama, M., Kajii, Y., Okuno, H., Tohyama, C., Bito, H., & Morris, R. G. M. (2011). Schema-Dependent Gene Activation and Memory Encoding in Neocortex. *Science*, *333*(6044), 891–895. https://doi.org/10.1126/science.1205274

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

Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: A review of evidence from a dual-process perspective. *Neuropsychologia*, *46*(7), 1787–1799. https://doi.org/10.1016/j.neuropsychologia.2008.01.004

Wagner, A. D., Maril, A., & Schacter, D. L. (2000). Interactions Between Forms of Memory: When Priming Hinders New Episodic Learning. *Journal of Cognitive Neuroscience*, *12*(Supplement 2), 52–60. https://doi.org/10.1162/089892900564064

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, *9*(9), 445–453. https://doi.org/10.1016/j.tics.2005.07.001

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

Wheelock, J. R., & Long, N. M. (2024). *The persistence of memory: Prior memory responses modulate behavior and brain state engagement* (p. 2024.04.05.588245). bioRxiv. https://doi.org/10.1101/2024.04.05.588245

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

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

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.

Zou, F., Wanjia, G., Allen, E. J., Wu, Y., Charest, I., Naselaris, T., Kay, K., Kuhl, B. A., Hutchinson, J. B., & DuBrow, S. (2023). Re-expression of CA1 and entorhinal activity patterns preserves temporal context memory at long timescales. *Nature Communications*, *14*(1), 4350. https://doi.org/10.1038/s41467-023-40100-8

CHAPTER III

RE-EXPRESSION OF CA1 AND ENTORHINAL ACTIVITY PATTERNS PRESERVES TEMPORAL CONTEXT MEMORY AT LONG TIMESCALES

From Zou, F., Wanjia, G., Allen, E. J., Wu, Y., Charest, I., Naselaris, T., Kay, K., Kuhl, B. A., Hutchinson, J. B., & DuBrow, S. (2023). Re-expression of CA1 and entorhinal activity patterns preserves temporal context memory at long timescales. Nature Communications, 14(1), 4350.

Abstract

Converging, cross-species evidence indicates that memory for time is supported by hippocampal area CA1 and entorhinal cortex. However, limited evidence characterizes how these regions preserve temporal memories over long timescales (e.g., months). At long timescales, memoranda may be encountered in multiple temporal contexts, potentially creating interference. Here, using 7T fMRI, we measured CA1 and entorhinal activity patterns as human participants viewed thousands of natural scene images distributed, and repeated, across many months. We show that memory for an image's original temporal context was predicted by the degree to which CA1/entorhinal activity patterns from the first encounter with an image were re-expressed during re-encounters occurring minutes to months later. Critically, temporal memory signals were dissociable from predictors of recognition confidence, which were carried by distinct medial temporal lobe expressions. These findings suggest that CA1 and entorhinal cortex preserve temporal memories across long timescales by coding for and reinstating temporal context information.

Introduction

Episodic memory fundamentally involves the ability to remember not only what happened in the past, but when it happened (Tulving, 2002). Indeed, placing memories in time critically enables experiences to be organized into personal narratives that span weeks, months, and years (Friedman, 2004). Yet, the majority of cognitive neuroscience studies of human memory only consider memory across relatively short timescales (overwhelmingly, within a single experimental session/day). At longer timescales, one of the particular challenges to retaining precise temporal memories is that previously-encoded information is likely to be 're-encountered' in new temporal contexts (Wixted, 2004). For example, remembering precisely when you first saw a particular movie may be complicated by re-watching that movie at a later date. Understanding how memories of specific temporal contexts are preserved when experiences are repeated over long timescales (days, weeks, months) requires identifying not only the neural structures that are involved, but the mechanistic contributions that these structures support.

Broadly, the medial temporal lobe (MTL) system is known to critically support episodic memory (Squire et al., 2004; Davachi, 2006; Eichenbaum et al., 2007). However, within the MTL system, hippocampal subfield CA1 and entorhinal cortex (ERC) have emerged as being particularly important for processing and remembering temporal information (Bellmund et al., 2020; Davachi & DuBrow, 2015; Eichenbaum, 2014; Lee et al., 2020; Ranganath & Hsieh, 2016). For example, so-called "time cells" in CA1 and "ramping cells" in ERC have been shown to code for elapsed time in rodents (Kraus et al., 2013; MacDonald et al., 2011; Pastalkova et al., 2008; Tsao et al., 2018), with similar effects recently observed in the human hippocampus and ERC (Reddy et al., 2021; Umbach et al., 2020). Importantly, although individual time cells typically operate over very short timescales (seconds), ensembles of time cells may provide temporal

context representations that span much longer timescales (Mau et al., 2018) and allow individual memories to be 'placed' in time (Eichenbaum, 2013). These temporal context representations are thought to integrate information about temporally adjacent events as well as ongoing internal operations or processes (Polyn et al., 2009). While human fMRI studies have provided important evidence that activation levels in the hippocampus and ERC are associated with the precision of temporal memory20,21, measures of activation, alone, are not able to capture the ensemble representations in which temporal context information is thought to be encoded (Howard, 2017; Polyn & Cutler, 2017).

Importantly, to the extent that CA1 and ERC do code for the temporal context in which events occur, there are multiple—and mechanistically distinct—ways in which these representations might preserve temporal memories. On the one hand, when a given stimulus is re-encountered in a new temporal context, CA1 and/or ERC may encode the new temporal context as distinct from the original context (O'Reilly & McClelland, 1994). Forming distinct temporal context representations across repeated encounters is potentially beneficial to temporal memory by improving discriminability of these contexts (Hsieh et al., 2014). On the other hand, when a stimulus is re-encountered in a new temporal context, this potentially creates an opportunity to reinstate the prior temporal context (Howard & Kahana, 2002; Polyn et al., 2009). For example, when a familiar movie is on television, this might trigger recall of the original temporal context in which the movie was encountered. Critically, this reinstatement should strengthen the association between the movie and its original temporal context (Sederberg et al., 2011). According to leading theoretical accounts, a stronger association between a given memory (e.g., the movie) and a particular temporal context (e.g., the movie's original temporal context) will directly support the ability to place that memory in time (Howard & Eichenbaum, 2013). Thus, in contrast to a context

distinctiveness account, a context reinstatement account makes the prediction that, when stimuli are re-encountered, memory for the original temporal context will be preserved to the extent that activity patterns in CA1 and/or ERC are similar to (or reinstate) the activity patterns expressed when the stimulus was first encountered.

Here we sought to characterize the neural mechanisms that preserve temporal context memory when events are re-encountered across long timescales (days to months). To address this, we describe a massive human fMRI experiment in which participants encountered thousands of natural scene images repeatedly during 30-40 scan sessions distributed over an 8-10 month window (Allen et al., 2022). After all scans were completed, participants performed a temporal memory task in which a subset of images were presented and participants were asked to estimate when each image was first encountered (on a scale that ranged from days to months in the past). The focus of our analyses was to test whether temporal memory precision was predicted by the degree to which patterns of neural activity expressed when images were first encountered were re-expressed when these images were re-encountered (a potential marker of context reinstatement). By leveraging the ultra-high field strength (7T) and high spatial resolution (1.8-mm) of our imaging protocol, we interrogated subregions of the hippocampus (including CA1) and surrounding MTL structures (including ERC). This experimental design yielded an unprecedented ability to understand how temporally-precise memories are preserved over long timescales that are critical for real-world memories.

Results

Precise temporal memory persists across months

Eight participants completed two experimental phases (Fig. 3.1a). The first phase consisted of a continuous recognition task conducted during fMRI scanning. The second phase consisted of a final memory test conducted outside of the scanner. During the continuous recognition phase, participants viewed 9,209-10,000 natural scene images across 30-40 fMRI sessions and indicated whether or not each image had previously been encountered at any point in the experiment (Fig. 3.1b). Each image was presented up to three times with these exposures pseudo-randomly distributed across the entire experiment (Fig. 3.1d). At least two days after completion of the last session of the continuous recognition phase, participants completed a final memory test on a subset of images (Fig. 3.1c). Each trial of the final memory test began with a recognition memory judgment on a 1-6 confidence scale (1: 'high confidence new', 6: 'high confidence old'). For images judged to be 'old', participants were also prompted to make frequency and temporal memory judgments. For the frequency judgment, participants were asked how many times they had seen the image during the continuous recognition phase (1, 2, 3, or 4 or more). For the temporal memory judgment, which is the primary focus of the present study, participants were instructed to position a marker along a continuous timeline when in the experiment each image was first encountered.

All participants performed above chance on the recognition memory test (Fig. 3.2a; hit rate greater than false alarm rate: $t_7 = 8.24$, $p < 0.001$, Cohen's $d = 1.56$, 95% confidence interval (CI) = [0.2, 0.36], two-tailed paired-sample t-test). Separating the data across three confidence levels (low, medium, and high) revealed that recognition memory accuracy (d') increased with

levels of subjective confidence (Fig. 3.2b; $F_{2,14} = 16.66$, $p < 0.001$, $\eta 2 = 0.70$, one-way repeated-measures ANOVA). Results for the frequency test are reported in Fig. 3.6.

Of critical interest was the accuracy of temporal memory judgments, which required participants to recall the first time each scene was encountered over the course of the up to 10-month experiment. To reduce the effects of non-linearity in temporal memory judgments (e.g., response bias towards the center of the timeline, see Methods and Fig. 3.7), we converted both the actual (objective) and the estimated (subjective) temporal positions to ranked positions for further analyses. Based on the ranks, we quantified item-wise temporal memory error by comparing the distance between the actual and estimated ranked positions (Fig. 3.1e). To determine temporal accuracy across participants, we ran a mixed-effects linear regression model for estimated against actual temporal position with participants as a random effect. Results from this analysis indicated that participants were able to place images in their correct temporal contexts with above-chance accuracy (Fig. 3.2c; group-level $\beta = 0.302$, $p < 0.001$, 95% CI = [0.24, 0.36]). We further evaluated temporal memory accuracy for each participant using a permutation test (see Methods). This analysis revealed that temporal memory performance was above chance for seven out of the eight participants (Fig. 3.2d; $ps < 0.01$; one participant: $p = 0.083$). The relatively high accuracy of the temporal memory judgments is notable when considering that participants were not informed that they would be tested on temporal memory until after all of the continuous recognition sessions.

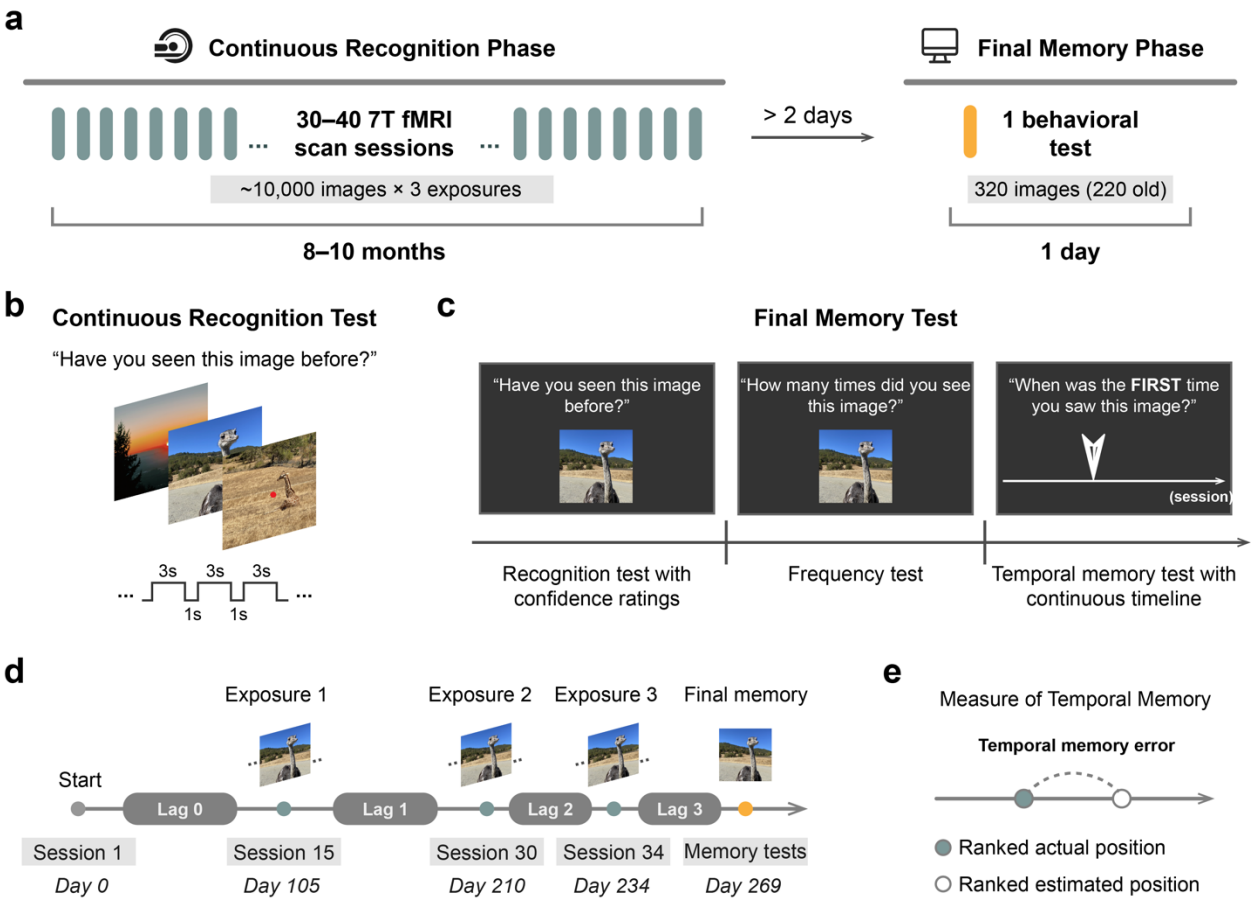

Figure 3.1. Experimental design. (a) Overview of experimental procedures: participants completed two experimental phases. The continuous recognition phase consisted of 30-40 separate fMRI scan sessions distributed across 8-10 months. Across these sessions, thousands of natural scene images were pseudo-randomly presented up to three times. After all of the scan sessions were completed, participants performed a final memory test on a subset of images outside of the scanner on a separate day (2-7 days later). **(b)** Continuous recognition test. While maintaining central fixation, participants viewed sequences of natural scenes and reported whether each image had been seen at any previous point in the experiment. **(c)** Final memory test. Each trial of the final memory test began with a recognition memory judgment in which participants made a recognition decision together with a confidence rating from 1-6 (1: 'high confidence new', 6: 'high confidence old'). For each image judged as 'old', a frequency test followed in which participants were asked how many times they had seen the image before (1, 2, 3, or 4 or more). Following that, participants were asked to indicate on a continuous timeline when the image in question was *first encountered* (temporal memory test; note this is a conceptual illustration of the task, see Methods for more information). **(d)** Timeline of an example image. Each old image used in the final memory test was presented three times during the continuous recognition phase and associated with four temporal lags. The first fMRI scan session of the continuous recognition phase for each participant corresponds to Day 0. All temporal lags were quantified in seconds and transformed with the natural logarithm for further analyses. **(e)** Behavioral measure of temporal memory. Item-wise

temporal memory error was quantified as the difference between the ranked actual and ranked estimated temporal positions.

CA1 and entorhinal representational similarity across exposures predicts temporal memory precision

The primary goal of the present study was to investigate whether the similarity (or dissimilarity) of MTL representations across repeated stimulus encounters predicts the accuracy of temporal memory judgments across long timescales. Accordingly, we examined the representational similarity between exposures of each of the images that were subsequently probed in the temporal memory test. Given our a priori interest in MTL structures, we focused on two manually segmented subfields of the hippocampus (CA1 and CA2/3/dentate gyrus, hereafter CA2/3/DG), along with ERC, perirhinal cortex (PRC), and parahippocampal cortex (PHC) (Fig. 3.3a). For each region of interest (ROI), we correlated the activity patterns between each pair of exposures of the same image (i.e., r(E1, E2), r(E2, E3), and r(E1, E3)). The resulting correlations were then Fisher-transformed prior to statistical testing. As a first step, we averaged across these pairwise correlations to generate a single similarity metric (across exposures) for each image (Fig. 3.3b). We then compared these similarity metrics for images associated with high versus low temporal memory precision (based on a participant-specific median split). Statistical significance of the difference between high and low temporal memory precision was evaluated using a permutation test that shuffles the images' temporal memory identities within each participant. Among the set of MTL ROIs, CA1 and ERC exhibited significantly greater pattern similarity across repeated exposures for high-precision images relative to low-precision images (Fig. 3.3c; CA1: $p = 0.004$; ERC: $p = 0.004$; permutation tests). The fact that temporal memory precision was associated with greater pattern similarity across exposures in CA1 and ERC is consistent with a

context reinstatement account, wherein the original temporal context is reinstated (and strengthened) during subsequent exposures.

We next performed several control analyses. First, because temporal memory precision increased as a function of the session position in which the first exposure occurred (recency effect, see Fig. 3.8), we repeated the analyses for CA1 and ERC while explicitly accounting for temporal lag information (Fig. 3.1d). Specifically, we ran a mixed-effects logistic regression model that predicted temporal memory precision from pattern similarity across exposures with temporal lags (lag 0-3) included as fixed effects and participant included as a random effect. This analysis confirmed that the relationship between pattern similarity in CA1/ERC and temporal memory precision remained significant when accounting for temporal lag information (Fig. 3.3d; CA1: β = 2.134, p = 0.005, 95% CI = [0.63, 3.64]; ERC: β = 3.207, p = 0.008, 95% CI = [0.83, 5.58]).

Second, we repeated the foregoing analyses for an early visual cortex ROI (V1) that would be sensitive to low-level visual information but would not be expected to code for temporal context. As expected, V1 pattern similarity across exposures did not differ for high- versus low-precision images (Fig. 3.3c; p = 0.25; permutation test) and was not a predictor of temporal memory precision (Fig. 3.3d; p = 0.376; logistic mixed-effects regression). Likewise, an additional, exploratory whole-brain analysis did not identify any cortical areas outside the MTL for which the relationship between pattern similarity and temporal memory was significant after correction for multiple comparisons (Table 3.1).

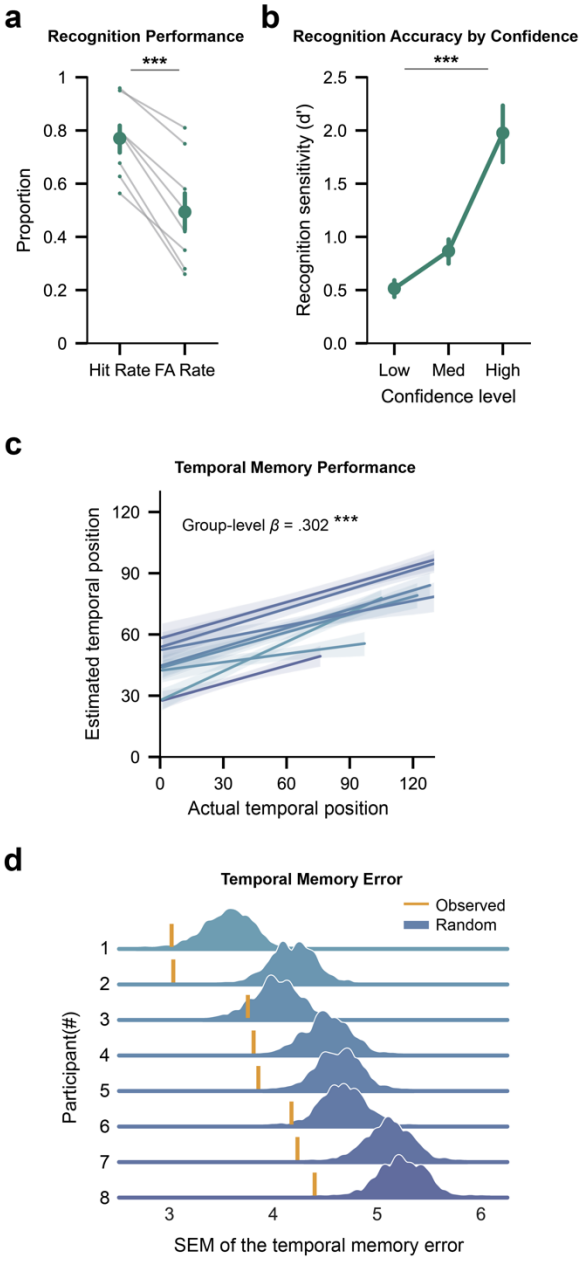

Figure 3.2. Behavioral results. (a) Recognition performance for each participant quantified by hit rate and false alarm (FA) rate. Hit rates were reliably above FA rates (two-tailed paired *t*-test; $t_7 = 8.24$, $p < 0.001$, Cohen's $d = 1.56$, 95% CI = [0.2, 0.36]), indicating above-chance recognition memory. **(b)** Overall recognition performance (d') separated by confidence levels. Recognition accuracy increased with subjective confidence levels (one-way repeated-measures ANOVA; $F_{2,14} = 16.66$, $p < 0.001$, $\eta^2 = 0.70$). **(c)** Correlation between estimated and actual temporal positions. Participants showed above-chance accuracy in temporal memory judgments (group-level $\beta = 0.302$, $p < 0.001$, 95% CI = [0.24, 0.36], linear mixed-effects regression, $n = 8$ independent participants). Each color shaded line indicates a participant. **(d)** Individual participant's temporal memory performance compared to chance level. Density plots compare the standard error of the

mean (SEM) of the observed temporal memory error (yellow line) to the null distribution (blue density; estimated by permuting estimated temporal judgments across images within each participant, $n = 1{,}000$ permutations). Throughout the figure, error bars reflect mean ± s.e.m.; dots or colors denote individual participants (n = 8); *** $p < 0.001$.

Third, and critically, we next tested whether the effects observed in CA1 and ERC were specific to temporal memory. To this end, we repeated the same mixed-effects regression model but now used recognition confidence as the dependent variable instead of temporal precision. Neither CA1 nor ERC exhibited significant relationships between pattern similarity and recognition confidence (ps > 0.10). In contrast, pattern similarity was a significant predictor of recognition confidence in PHC (Fig. 3.3e; $\beta = 0.799$, $p < 0.001$, 95% CI = [0.34, 1.25]). A follow-up control analysis which included recognition confidence together with pattern similarity as fixed effects in a mixed-effects regression model confirmed that pattern similarity in CA1 and ERC predicted temporal memory precision when accounting for recognition confidence (CA1: $\beta = 2.073$, $p = 0.007$, 95% CI = [0.56, 3.58]; ERC: $\beta = 3.132$, $p = 0.010$, 95% CI = [0.75, 5.51]). These results help constrain accounts of why pattern similarity in CA1/ERC predicted temporal memory precision. Namely, they argue against the possibility that the relationships between CA1/ERC pattern similarity and temporal memory precision were a secondary consequence of overall memory strength for the images. Rather, pattern similarity across exposures in CA1 and ERC specifically predicted better memory for when (in time) images were first encountered.

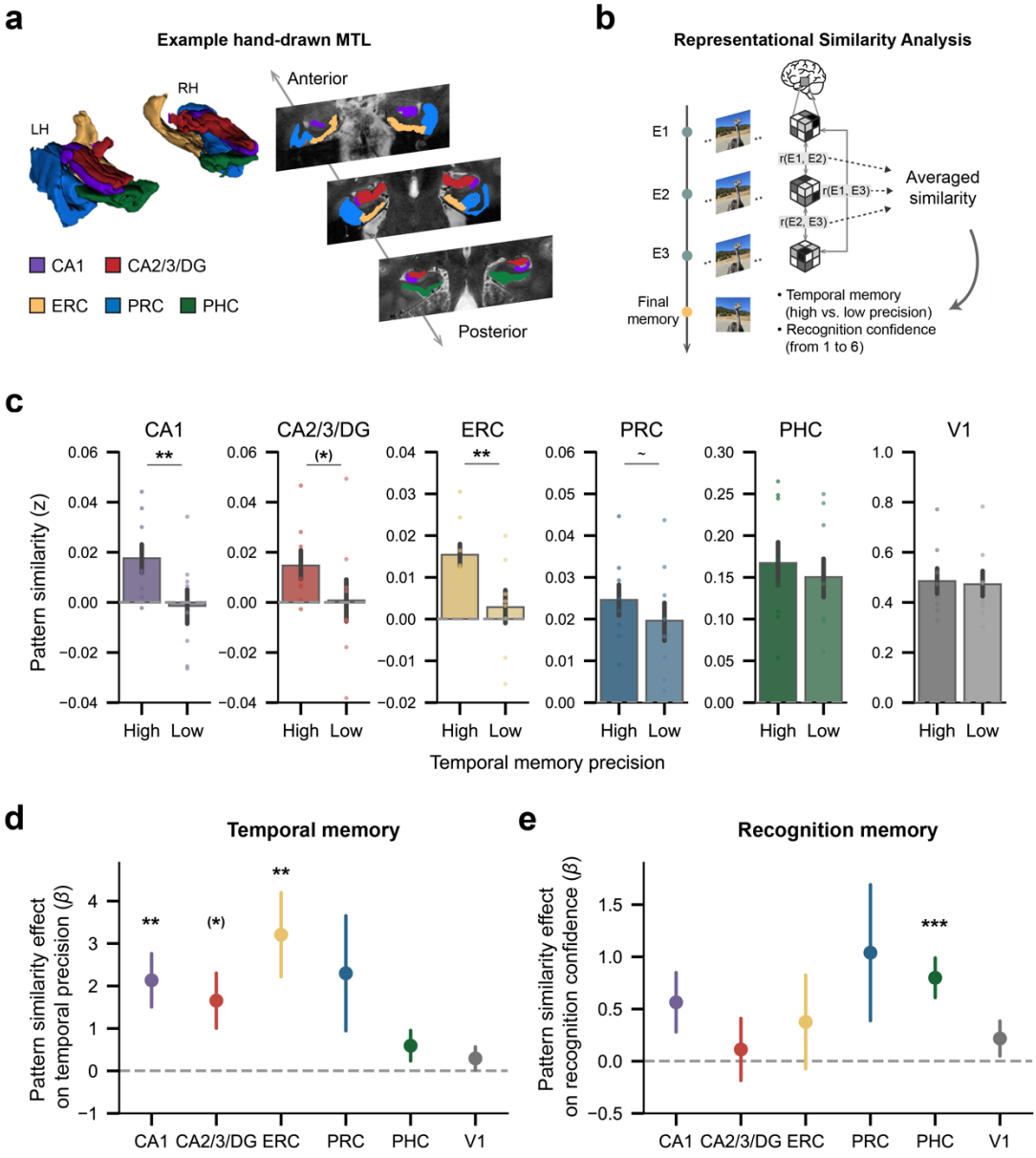

Figure 3.3. CA1 and entorhinal representational similarity predicted temporal memory precision, but not recognition confidence. (a) Manually drawn ROIs for MTL subregions of an example participant: CA1 (purple), CA2/3/DG (red), ERC (yellow), PRC (blue), and PHC (green). LH/RH: left/right hemisphere. **(b)** Schematic depiction of representational similarity analysis. **(c)** Pattern similarity difference between high- and low-precision images (median split) across MTL subregions and a control early visual region (V1). CA1 and ERC showed greater pattern similarity across exposures for high-precision images relative to low-precision images (CA1: p = 0.004; ERC: p = 0.004; one-sided permutation tests, n = 1,000). CA2/3/DG showed similar effect but did not survive correction for multiple comparisons (*p(uncorrected)* = 0.023). **(d)** Relationship between pattern similarity across exposures and temporal memory precision. Pattern similarity across repeated exposures in CA1 and ERC predicted temporal memory precision (high vs. low)

while accounting for temporal lag information (CA1: $\beta = 2.134$, $p = 0.005$, 95% CI = [0.63, 3.64]; ERC: $\beta = 3.207$, $p = 0.008$, 95% CI = [0.83, 5.58]; logistic mixed-effects regression, $n = 8$ independent participants). A similar effect was also observed in CA2/3/DG (*p(uncorrected)* = 0.037), but did not survive correction for multiple comparisons. **(e)** Relationship between pattern similarity across exposures and recognition confidence. Pattern similarity across repeated exposures in PHC predicted recognition confidence while accounting for temporal lag information ($\beta = 0.799$, $p < 0.001$, 95% CI = [0.34, 1.25]; liner mixed-effects regression). Throughout the figure, error bars reflect mean ± s.e.m.; dots denote independent participants ($n = 8$); ~$p < 0.10$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. Parentheses indicate ROIs that did not survive multiple comparison correction.

Similarity between first and second exposures uniquely predicts temporal memory

Having demonstrated that CA1 and ERC pattern similarity across repeated exposures predicts temporal memory for an image's first exposure, we next sought to determine which pair of image exposures was most predictive of temporal memory. From a context reinstatement perspective, similarity between the first exposure (E1) and the second exposure (E2) should be uniquely important because E2 provides the first opportunity to reinstate the temporal context from E1. To test this, we first compared pattern similarity for high- and low-precision images for each pair of image exposures (E1-E2, E2-E3, and E1-E3). Statistical significance of the difference between high- and low- precision images for each exposure pair was computed a permutation analysis in which, for each participant and exposure pair, we randomly shuffled the images' temporal memory precision labels. For both CA1 and ERC, E1-E2 similarity was significantly greater for high- than low-precision images (Fig. 3.4a; CA1: p = 0.015; ERC: p = 0.007; permutation tests). However, both regions also exhibited similar effects for E2-E3 similarity (Fig. 3.4a; CA1: p = 0.025; ERC: p = 0.036, permutation tests). Neither region exhibited a significant effect for E1-E3 similarity (Fig. 3.4a; ps > 0.28).

To further explore this pattern of results, we performed three follow-up sets of analyses. First, in order to control for potential temporal lag effects (Fig. 3.8), we ran a mixed-effects logistic

regression model that predicted temporal memory from pattern similarity of each exposure pair (E1-E2, E2-E3 and E1-E3 as separate dependent variables in one regression model) while including lag information. For both CA1 and ERC, E1-E2 similarity significantly predicted temporal memory (Fig. 3.4c; CA1: β = 1.048, p = 0.014, 95% CI = [0.21, 1.88]; ERC: β = 1.565, p = 0.022, 95% CI = [0.22, 2.91]). Effects were marginally significant for E2-E3 similarity (ps $<$ 0.10), and not significant for E1-E3 similarity (ps > 0.68).

Second, in order to more directly assess whether E1-E2 similarity contained predictive power above and beyond that of other exposure pairs, we compared the performance of several models that did or did not include various exposure pairs. That is, we tested whether model performance was significantly improved when E1-E2 similarity was added to models that only included E2-E3 and E1-E3 similarity. For both CA1 and ERC, adding E1-E2 as a predictor significantly improved the model's performance (CA1: χ^2 = 6.147, p = 0.013; ERC: χ^2 = 5.315, p = 0.021). In contrast, adding E2-E3 and E1-E3 similarity as predictors to models with just E1-E2 similarity did not improve the model's performance (ps > 0.15). These results established that E1-E2 similarity was uniquely important for subsequent temporal memory judgments, as would be predicted by a context reinstatement account.

Third, to further establish whether E1-E2 similarity was uniquely important for temporal memory, we tested for relationships between different exposure pairs and recognition memory confidence. Interestingly, although PHC pattern similarity across exposures was highly predictive of subsequent recognition memory confidence (Fig. 3.3e), this effect was not driven by E1-E2 similarity (Fig. 3.4b; p = 0.482, permutation test; Fig. 3.4d; p = 0.225; linear mixed-effects regression). Instead, E1-E3 similarity in PHC significantly predicted recognition confidence (Fig. 3.4b; p = 0.002; Fig. 3.4d; β = 0.473, p = 0.018, 95% CI = [0.08, 0.86]). Along with the results

described above, these findings provide a qualitative dissociation between the predictors of temporal memory versus recognition memory. That is, whereas pattern similarity between the first and second exposure of an image was uniquely important for remembering when the image was first encountered, it was relatively less important for recognizing whether the image was previously encountered.

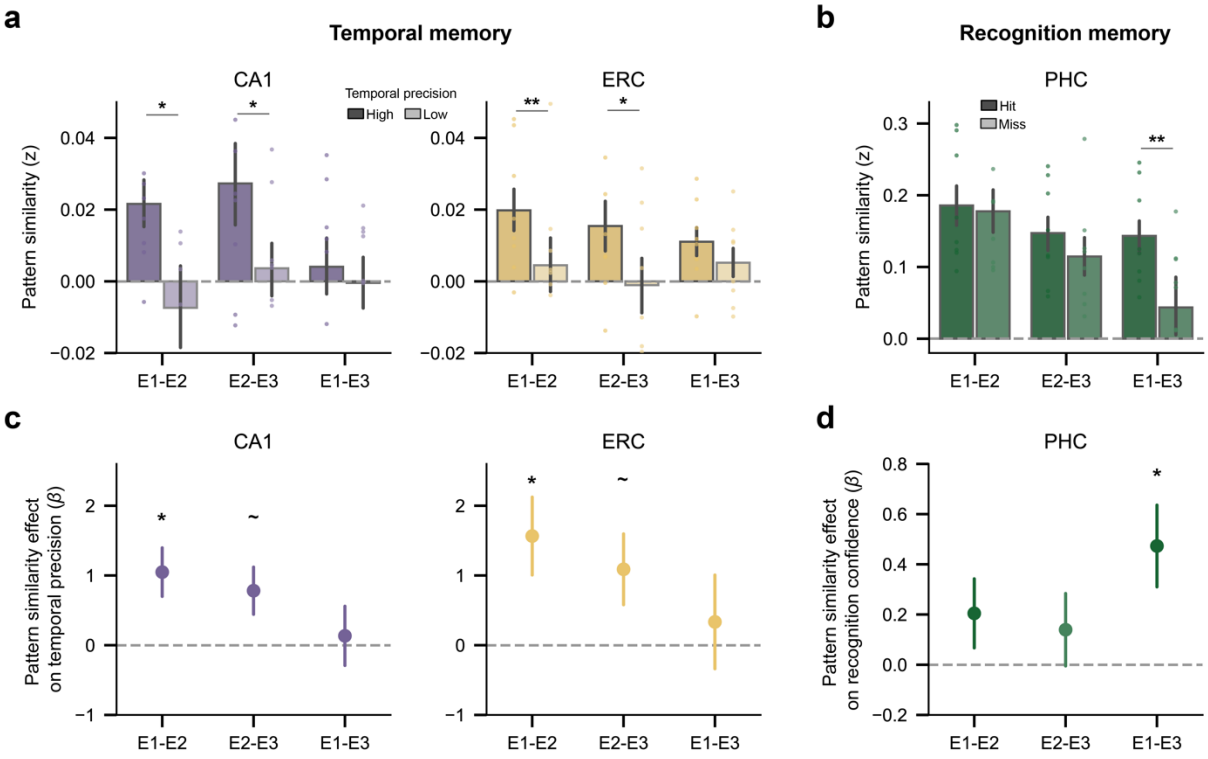

Figure 3.4. Pattern similarity between the first and second exposure in CA1 and ERC was uniquely important for temporal memory. (a) CA1/ERC pattern similarity between high- and low-precision images for each pair of image exposures. CA1 and ERC showed greater pattern similarity for high-precision images relative to low-precision images in E1-E2 (CA1: p = 0.015; ERC: p = 0.007; permutation test, n = 1,000) and E2-E2 (CA1: p = 0.025; ERC: p = 0.036; permutation test). **(b)** PHC pattern similarity between hits and misses in recognition memory for each pair of image exposures. PHC showed greater pattern similarity for hits relative to misses in E1-E3 (p = 0.002; permutation test, n = 1,000). **(c)** Relationship between pattern similarity for each pair of image exposures in CA1/ERC and temporal memory precision while accounting for temporal lag information. For both CA1 and ERC, E1-E2 pattern similarity was significantly predictive of temporal memory precision (CA1: β = 1.048, p = 0.014, 95% CI = [0.21, 1.88]; ERC: β = 1.565, p = 0.022, 95% CI = [0.22, 2.91]; logistic mixed-effects regression, n = 8 independent

participants). **(d)** Relationship between pattern similarity for each pair of image exposures in PHC and recognition confidence while accounting for temporal lag information. Recognition confidence was predicted by E1-E3 pattern similarity in PHC (β = 0.473, p = 0.018, 95% CI = [0.08, 0.86]; linear mixed-effects regression). Error bars reflect mean ± s.e.m.; dots denote independent participants (n = 8); ~p < 0.10; *p < 0.05; **p < 0.01.

CA1 and ERC predict temporal memory via image-specific representations

While all of the preceding representational similarity analyses were performed by correlating activity patterns across repeated exposures of the same stimulus (i.e., image-specific correlations), these analyses do not guarantee that the information that predicted temporal memory precision was specific to individual images. Namely, it is possible that temporal memory precision benefited from generic memory processes or attentional states that generalized across images (e.g., states optimized for memory encoding (Ezzyat et al., 2017)). While this possibility would still support a role for CA1 and ERC in encoding temporal information, a temporal context reinstatement account fundamentally predicts reinstatement of the specific temporal context in which an image was encoded.

To assess whether temporal memory was predicted by image-specific pattern similarity, we conducted two additional analyses (restricted to E1-E2 similarity). First, for all of the images tested in the temporal memory test, we permuted the E1-E2 mappings by shuffling images' E2 within each participant. We then calculated the resulting E1-E2 pattern similarity scores and a corresponding distribution of beta values reflecting the relationships with temporal memory (see Methods for details). Critically, for both CA1 and ERC, the relationship between 'intact' E1-E2 similarity and temporal memory was significantly stronger (higher beta values) than the permuted values (Fig. 3.5a; CA1: p = 0.019; ERC: p = 0.025). These data provide important evidence that temporal memory precision was predicted by image-specific pattern similarity in CA1 and ERC.

As a follow-up to the preceding analysis, we ran a final analysis to address whether apparent image-specific effects might be due to general memory states and/or differences in coarse temporal context information (i.e., session effects). Thus, for each image included in the temporal memory test (a 'target'), we identified control images ('foils') such that the targets and foils shared the same E1 session number, but not scanning run (to avoid potential contamination from autocorrelation in the fMRI data), and the same E2 session number (but not run; Fig. 3.5b). To match recognition memory with targets, foils were only included in this analysis if they were correctly rejected at E1 and successfully recognized at E2 and E3 (see Methods for details). This allowed us to compute similarity between target E1 and target E2 (target similarity) and target E1 and foils E2 (foil similarity). The difference between these measures (target similarity – foil similarity) was then used as a predictor of temporal memory precision. Indeed, this similarity difference score significantly predicted temporal memory precision for both CA1 (Fig. 3.5c; β = 0.864, p = 0.033, 95% CI = [0.07, 1.66]) and ERC (Fig. 3.5c; β = 1.308, p = 0.047, 95% CI = [-0.02, 2.60]). These findings lend further support to the idea that temporal memory precision was related to image-specific pattern similarity measures and specifically argue against potential confounds due to generic memory-related processes or session effects. The fact that these effects held when carefully controlling for session effects is notable because it provides evidence against the possibility that pattern similarity only captured coarse-level temporal context (session information). Rather, to the extent that the pattern similarity measure captured temporal context information, these findings suggest a relatively 'local' temporal context representation that differentiated between images within the same session (day).

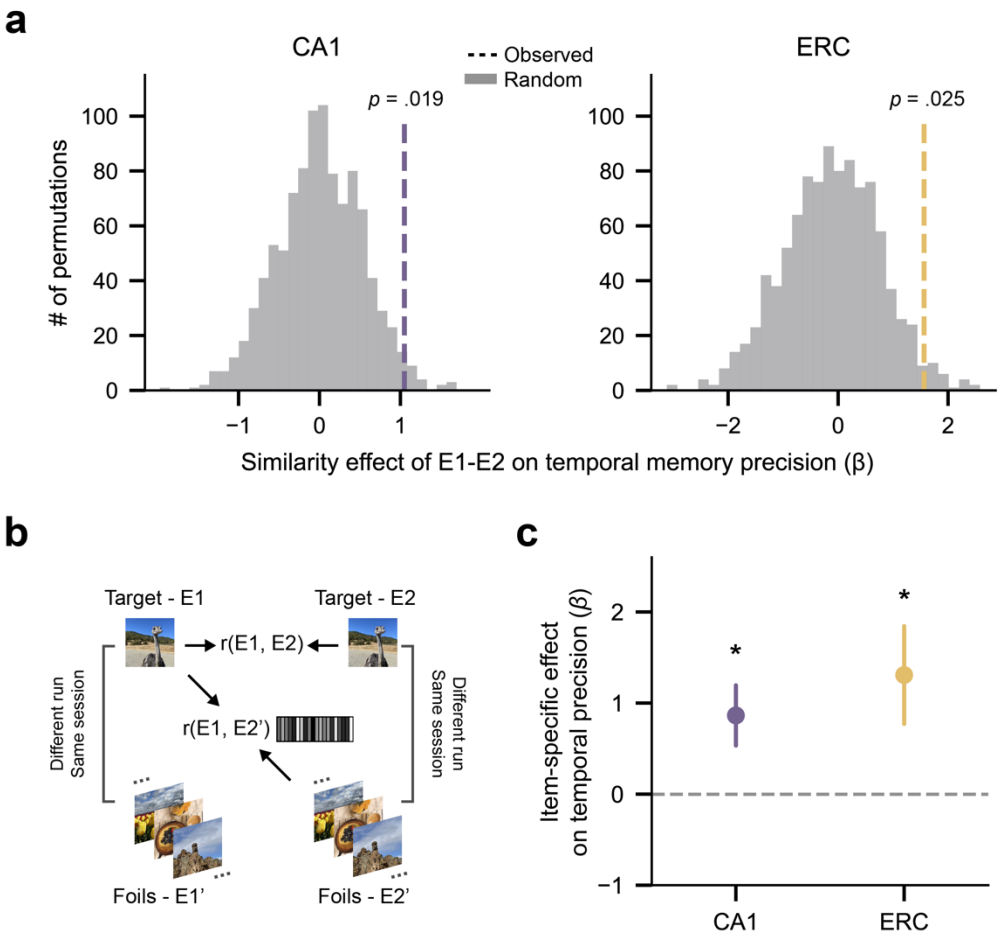

Figure 3.5. Representational image-specificity analyses. (a) Intact compared to permuted similarity effect. E1-E2 pattern similarity compared to permuted similarity exhibited a stronger effect on temporal memory precision in both CA1 and ERC (CA1: $p = 0.019$; ERC: $p = 0.025$; permutation tests, $n = 1{,}000$). **(b)** Schematic illustration showing how target similarity and foil similarity were computed for an example image (see Methods for details). **(c)** Relationship between image-specific pattern similarity (target similarity – foil similarity) in CA1/ERC and temporal memory precision. Image-specific pattern similarity was significantly predictive of temporal memory precision in both CA1 and ERC (CA1: $\beta = 0.864$, $p = 0.033$, 95% CI = [0.07, 1.66]; ERC: $\beta = 1.308$, $p = 0.047$, 95% CI = [-0.02, 2.60]; logistic mixed-effects regression, $n = 8$ independent participants). Error bars reflect mean ± s.e.m.; $\sim p < 0.10$; $^*p < 0.05$.

Discussion

The ability to remember when events occurred in time is fundamental to human experience. However, retaining precise temporal memories is complicated by the fact that real-world episodic memories span long timescales (days, weeks, months and beyond) and by the fact that events may

re-occur in multiple contexts over those long timescales (e.g., a movie you have viewed several times over the past year). To date, there is remarkably little evidence characterizing how the human brain preserves temporal memories in the face of these challenges. Here, we show that when events re-occur over long timescales (at lags up to several months), the re-expression of distributed, event-specific activity patterns in CA1 and ERC preserves memory for the original temporal context of an event (i.e., memory for when an event first occurred). These findings are consistent with and bridge between prior human and rodent studies implicating CA1 and ERC in temporal processing and temporal memory. However, our findings also go beyond existing evidence by providing a mechanistic account of how CA1 and ERC preserve temporal memories and demonstrating these relationships at uniquely long timescales.

While there is a rich history characterizing temporal memory in human behavioral and neuroimaging studies (Friedman, 1993; Wang & Diana, 2017), it is striking how few of these studies have considered temporal memory across timescales that exceed a single experimental session. Indeed, our approach of testing temporal memory for images that were distributed across dozens of experimental sessions/scans spanning 8-10 months is unprecedented. Considering that the overwhelming majority of real-world episodic memories span days, weeks, months and years, it is imperative to understand the neural mechanisms that support temporal memory at these timescales. Although it is intuitively obvious that humans can and do retain temporal memories over long timescales, it is nonetheless remarkable that participants in the current study were generally successful at recalling the initial temporal context for images presented at the final memory test given that (a) these images were drawn from a pool of tens of thousands of images, (b) the delay between the initial exposure and the final memory test ranged from days to almost a year, and (c) each image was presented in multiple temporal contexts, creating potential

interference. Thus, by simulating the challenges that are inherent to real-world temporal memory, our experimental paradigm provides a unique opportunity to characterize the underlying neural mechanisms.

By leveraging representation-based analyses to track patterns of activity across repeated stimulus exposures and distinct temporal contexts, we were able to gain critical insight into the mechanisms through which CA1 and ERC contribute to temporal memory. In particular, our findings strongly align with a context reinstatement account. According to temporal context models (Howard & Kahana, 2002; Polyn et al., 2009), context representations—reflected in distributed patterns of neural activity—gradually change over time and are reinstated when a stimulus is subsequently remembered (DuBrow & Davachi, 2014; Folkerts et al., 2018; Howard et al., 2012; Kragel et al., 2021; Manning et al., 2011; Manns et al., 2007; Yaffe et al., 2014). This temporal context reinstatement is thought to directly support the ability to remember—or infer—when in time stimuli were previously encountered (Howard & Eichenbaum, 2013). From this perspective, our finding that greater pattern similarity across exposures preserved memory for an event's original temporal context can be explained in terms of the original context representation (elicited during E1) being reinstated and strengthened during subsequent exposures (E2, E3) (Ritchey et al., 2013) and ultimately guiding temporal memory judgments at the end of the experiment. In fact, this account also readily explains our finding that similarity between the first and second exposure (E1, E2) was uniquely important for temporal memory. Namely, E2 represented the first potential 'reminder' of E1's temporal context. Interestingly, although we tested for re-expression of E1's activity patterns by explicitly re-exposing participants to the same stimulus multiple times (E2, E3), our findings likely generalize to situations where stimuli are not explicitly re-exposed (or re-encountered). Indeed, human neuroimaging studies (unrelated to

temporal memory) have found that offline, spontaneous reinstatement of episodic memories not only occurs, but it strengthens memories in much the same way that online, cued reinstatement does (Tambini & Davachi, 2019).

While a context reinstatement account makes a clear prediction that better memory for the original temporal context should be associated with greater representational similarity across exposures, it is notable that a memory interference account (O'Reilly & McClelland, 1994) suggests an entirely opposite prediction: that temporal memory would benefit from greater contextual distinctiveness across exposures (i.e., less similarity). More specifically, greater contextual distinctiveness would putatively be expected to reduce interference between the various temporal contexts in which an event occurred (E1, E2, E3). That said, there are several examples in the memory interference literature where reinstatement of prior experiences during new learning can, in fact, protect memories from interference (Chanales et al., 2019; Kuhl et al., 2010). Moreover, it is important to note that a context reinstatement account for CA1 and ERC does not exclude the possibility that other MTL regions (e.g., CA3) might simultaneously contribute to temporal memory by emphasizing differences between temporal contexts (Copara et al., 2014; Dimsdale-Zucker et al., 2018; Schapiro et al., 2012; Yassa & Stark, 2011).

The fact that we specifically identified CA1 and ERC as being important for temporal memory at long timescales—and the implication that these regions supported temporal context reinstatement—is striking in light of accumulating evidence documenting time sensitive cells within rodent CA1 and ERC (MacDonald et al., 2011; Pastalkova et al., 2008; Tsao et al., 2018). It has been speculated that ensembles of these cells allow for the coding of gradually-drifting temporal context representations which become bound to individual events (Eichenbaum, 2013) and reinstated when events are remembered (Howard & Kahana, 2002; Polyn et al., 2009).

Importantly, although individual cells may only operate across very short timescales (seconds), ensemble representations can track temporal information over much longer timescales—from minutes to days (Mankin et al., 2012; Mau et al., 2018). Here, we were not able to directly measure or identify time cells because, even with the relatively high spatial resolution of the current fMRI data, each voxel likely pooled across tens or hundreds of thousands of cells. However, the representation-based analyses we employed are potentially well-suited to capturing gradual changes in ensemble-level context representations (Bellmund et al., 2022; Deuker et al., 2016; Ezzyat & Davachi, 2014; Howard et al., 2012; Hsieh et al., 2014; Lositsky et al., 2016; Manning et al., 2011). In contrast, although several prior studies of human memory have also implicated CA1 and ERC in memory for when events occurred (Jenkins & Ranganath, 2010; Montchal et al., 2019), most of these studies have not employed representation-based analyses and, therefore, are not amenable to testing or capturing temporal context representations. Thus, our approach and findings uniquely bridge between evidence of time cells in rodents, theoretical models of temporal context, and prior studies of temporal memory in humans.

An additional essential consideration in understanding neural mechanisms that specifically relate to temporal memory is to establish that any apparent effects related to temporal memory were not derivative from more general effects of memory strength. Specifically, as memories decay over time, temporal judgments could potentially be inferred from the strength of memories themselves (Hinrichs, 1970; Hintzman, 2005; Jenkins & Ranganath, 2016). This is of particular concern given the very long timescales involved in the current study. However, several theoretical perspectives propose that memory for time is dissociable from memory strength (DuBrow & Davachi, 2017; Friedman, 1993; Howard et al., 2015). Here, our final memory test separately measured recognition confidence (a proxy for overall memory strength) and temporal memory,

allowing us to conduct several targeted analyses aimed at teasing apart these two expressions of memory. First, we found that the relationships between CA1/ERC and temporal memory precision remained significant in a regression model that included recognition confidence as a covariate. Second, consistent with prior arguments that distinct MTL subregions are involved in 'item-based' versus 'context-based' memory (Eichenbaum et al., 2012), we found that pattern similarity measures in PHC predicted recognition confidence but not temporal memory, whereas pattern similarity measures in CA1 and ERC predicted temporal memory but not recognition confidence. Finally, when considering pattern similarity across specific pairs of image exposures, temporal memory (defined here as memory for when the first exposure occurred) was best predicted by pattern similarity between the first and second exposures, consistent with a context reinstatement account. In contrast, recognition confidence was best predicted by pattern similarity between the first and third exposures, potentially indicating that the last (third) exposure was relatively more influential to memory strength (also see Fig. 3.8). Together, these data points provide important, converging evidence that temporal memory judgments in the current study were not derived from the overall memory strength. More generally, our findings reinforce theoretical accounts that emphasize the distinction between memory for 'when' an event occurred versus 'whether' an event occurred (Davachi, 2006; Diana et al., 2007; Eichenbaum et al., 2007; Ranganath, 2010).

In conclusion, here we show that memory for the temporal context in which an event initially occurred is preserved via the re-expression of activity patterns in human CA1 and ERC. Critically, we show that these dynamics operate across—and support memory at—long timescales (from days to months). These findings complement yet significantly advance existing evidence from rodents and humans implicating the hippocampal-entorhinal system in representing and remembering time. In particular, our findings suggest that distributed patterns of activity in CA1

and ERC encode and reinstate temporal context information, thereby preserving memory for when events occurred.

Methods

Participants

Eight participants took part in the study (two self-identified males, six self-identified females; age range: 19–32). All participants were right-handed with no known cognitive deficits nor color blindness and with normal or corrected-to-normal vision. Participants were naïve to the experimental manipulation and were not involved in the design nor planning of the study. Informed written consent was obtained from all participants before the start of the study, and the experimental protocol was approved by the University of Minnesota Institutional Review Board.

Design and procedure

Data used in this study were collected as part of the Natural Scenes Dataset (NSD; http://naturalscenesdataset.org), and included two parts: a continuous recognition phase conducted in the fMRI scanner and a behavioral final memory phase (Fig. 3.1a).

Continuous recognition phase.

A detailed description of the continuous recognition phase has been reported in a previous publication (Allen et al., 2022). Briefly, for each participant, the continuous recognition phase was split across 40 scan sessions in which 10,000 distinct color natural scenes would be presented three times spaced pseudo-randomly over the course of all scan sessions using Psychophysics Toolbox 3.0.14. Each scan session consisted of 12 runs (750 trials). Distributions of image presentations were controlled such that both short-term and long-term re-exposures were probed (see Stimuli

section below). Four of the participants completed the full set of 40 NSD scan sessions. Due to constraints on participant and scanner availability, each of the other four participants completed 30-32 scan sessions. In these collected data, each participant viewed 9,209-10,000 distinct images and participated in 22,500-30,000 trials. Each trial lasted 4 s, consisting of the presentation of an image for 3 s and a following 1-s gap. Participants were instructed to perform a continuous recognition task in which they reported whether the current image had been seen at any previous point in the experiment.

Final memory phase.

At least two days (range: 2-7 days) after completion of the continuous recognition phase, a final memory test was administered outside of the scanner. Participants were not informed about the final memory test in advance. During the final memory phase, participants viewed a subset of old images (220 per participant) from the continuous recognition phase randomly intermixed with novel images (100 per participant) and completed different types of memory probes. The final memory phase consisted of 320 trials, with up to three judgements per trial. Each trial began with a recognition test in which participants performed an old or new judgment with a confidence rating on a scale of 1 to 6 (1: 'high confidence new', 2: 'medium confidence new', 3: 'low confidence new', 4: 'low confidence old', 5: 'medium confidence old', 6: 'high confidence old'). For images judged as "old", a frequency test followed in which participants were asked to indicate how many times they had seen each image (1, 2, 3, or 4 or more times). Following the frequency test, participants performed a temporal memory test using a timeline. In this test, participants were asked to indicate, on a continuous timeline with tick marks to represent each session, when in the experiment they thought each image was first encountered (Fig. 3.1c, right). The length and labels of the timeline vary across participants, depending on how many sessions they completed in the

continuous recognition phase. Participants were encouraged to use the full length of the scale, with the left endpoint representing the beginning of the continuous recognition phase and the right endpoint representing the end. Participants used a cone to mark the temporal location on the line and were instructed to indicate their confidence in response via adjusting the size of the cone, with smaller cones representing higher confidence and bigger cones representing lower confidence. Given the primary focus of the present study concerns temporal memory precision, we only analyzed the estimates of temporal location as illustrated in Fig. 3.1c. All tests in the final memory phase were self-paced with a timeout of 30 s.

Stimuli

All images used in this study were taken from the Microsoft Common Objects in Context (COCO) database (Lin et al., 2014).

Continuous recognition phase.

For the continuous recognition phase, a total of 73,000 images were prepared with the intention that each participant would view 10,000 distinct images (9,000 unique images and 1,000 shared images across participants) three times each over the course of 40 scan sessions. To prevent the recognition task from becoming too difficult (and risking loss of morale), each image was randomly placed three times on a circle according to a probability distribution created by mixing a relatively narrow von Mises distribution and a uniform distribution. Across all scan sessions, the mean number of distinct images shown once, twice, and all three times within a typical session is 437, 106, and 34, respectively.

Final memory phase.

The final memory phase included a total of 320 images for each participant: 220 old images and 100 novel images. Old images were selected from the continuous recognition phase that participants completed during fMRI scanning with each image having been presented three times during the continuous recognition phase. Novel images were from the COCO dataset but were not presented during the continuous recognition phase.

The 220 old images included in the final memory test were comprised of two sets of images selected according to different criteria. The first set comprised 120 images that were selected based on behavioral accuracy and the distribution of exposures during the continuous recognition phase. Specifically, these images were selected based on the following three criteria: (1) Each of the 120 images was associated with a correct behavioral response at each exposure during the continuous recognition phase—i.e., correct rejection (E1), hit (E2), and hit (E3). (2) To promote overall temporal memory performance, approximately half of the selected images were associated with a first exposure (E1) that occurred during the last eight scan sessions (for one participant, this was adjusted to the last ten scan sessions in order to have enough trials given their performance in the continuous recognition phase); the other half of the selected images were associated with a first exposure that occurred from the rest of the scan sessions (i.e., earlier sessions). (3) For each half of the images, there was an additional constraint on the spacing between exposures, with one-third of the images having all three exposures within one scan session, one-third with the last two exposures in the same session, and the rest either with the first two exposures in the same session or with all three exposures across different sessions.

The second set of 100 old images included during the final memory test were selected to maximally span semantic space (see the NSD data paper (Allen et al., 2022) for details), without consideration of behavioral accuracy and distribution of exposures. Briefly, this was done by

computing shifted inverse frequency sentence embeddings for the sentence captions, and using a greedy approach to determine the subset of 100 images that maximize the average distance between each image's embedding and its closest neighbor. The motivation for sampling semantic space was unrelated to the goals of the current manuscript. However, in order to increase statistical power, we opted to include images from the second set in our analyses but only for images that were associated with correct behavioral responses for each of the three exposures in the continuous recognition phase (as was the case for the first set). This resulted in a total of 143-170 images that were used for analyses for each participant (see Fig. 3.9 for the distribution of images' first exposures across participants).

MRI data acquisition and preprocessing

The imaging data was collected as part of the NSD at the Center for Magnetic Resonance Research at the University of Minnesota. In brief, functional data and a few additional anatomical measures were collected using a 7T Siemens Magnetom passively-shielded scanner with a single-channel-transmit, 32-channel-receive RF head coil (Nova Medical, Wilmington, MA). Functional data was acquired using whole-brain gradient-echo echo-planar imaging (EPI) at 1.8-mm resolution and 1.6-s repetition time. In addition to the EPI scans, for the purposes of hippocampal segmentation, a high-resolution T_2-weighted scan was acquired during one of the 7T scan sessions. T_1- and T_2-weighted structural scans were collected using a combination of a 3T Siemens Prisma scanner and a standard Siemens 32-channel RF head coil.

Functional data were pre-processed by performing one temporal resampling to correct for slice time differences and one spatial resampling to correct for head motion within and across scan sessions, EPI distortion and gradient non-linearities. Informed by the original data paper (Allen et

al., 2022), the current study used the upsampled 1.0-mm high-resolution preparation of the NSD data in order to optimally partition the functional data into regions of interest defined using high-resolution anatomical images.

Parameter estimates (beta weights) reflecting fMRI response amplitudes evoked by each trial were estimated using a general linear model (GLM) approach as described in the NSD data paper (Allen et al., 2022). Notably, our approach (which corresponds to "beta version 2" in the NSD data paper) involved generating voxel-specific hemodynamic response functions (HRFs). Briefly, the pre-processed time-series data were fitted multiple times with a single-trial GLM, each time using a different HRF from a library of HRFs. For each voxel, the HRF that provided the best fit to the data was identified and single-trial betas were then generated using that HRF. Betas were then converted to units of percent BOLD signal change by dividing amplitudes by the mean signal intensity observed at each voxel and multiplying by 100. Our decision to use voxel-specific HRFs was made a priori and was motivated by evidence from the original data paper that it leads to fewer artifacts.

Regions of interest (ROIs)

The medial temporal lobe (MTL) ROIs were manually drawn on the high-resolution T_2 images obtained for each participant, following a 7T protocol for segmentation of MTL subregions (Berron et al., 2017). Labels were defined on the raw high-resolution T_2 volume, and were mapped via an affine transformation to subject-native anatomical space. The MTL ROIs included bilateral CA1, CA2/3/dentate gyrus, entorhinal cortex (ERC), perirhinal cortex (PRC), and parahippocampal cortex (PHC). Example MTL ROIs from one participant were depicted in Fig. 3.3a. We also included the primary visual cortex (V1) as a control region. The bilateral V1 ROI

was manually drawn on cortical surfaces based on results of a population receptive field experiment from the NSD, and were then mapped to volumetric format. Cortical ROIs for the whole-brain parcel level analysis were defined by a multi-modal cortical parcellation from the Human Connectome Project (Glasser et al., 2016).

Behavioral data analyses

Overall performance for the temporal memory test was quantified by regressing each participant's subjective estimate of when an image was first encountered against the actual (objective) time (Figure 3.2c). Note that there is a general response bias among participants toward the center of the timeline ("raw estimated position", see Figure 3.6). To account for this response bias and potential non-linearity, the estimated and actual temporal positions used in all analyses in the current paper were converted to ranks according to each individual's marked positions on the timeline and the actual temporal positions in the continuous recognition phase, respectively. To quantify item-wise temporal memory error, we calculated the absolute difference between the ranked estimated temporal position and the ranked actual position (Figure 3.1e). To test whether each participant had above-chance temporal memory performance, we compared the observed temporal memory error against a null distribution of permutations (1,000 iterations), in which the subjective estimates were randomly shuffled across trials for each participant and the temporal memory error was recomputed for each iteration. Due to the non-normal distribution of temporal error (absolute value of the difference of two rank distributions, see Figure 3.10), we divided temporal memory trials into 'high-precision' and 'low-precision' groups by performing a median split of temporal error for each participant. This was an a priori decision (i.e., made before conducting any fMRI analyses).

To control for temporal lag information and test for relationships between lag and subsequent memory performance (Figure 3.8), as illustrated in Figure 3.1d, four temporal lags were calculated for each image: the lag between the beginning of the continuous recognition phase and the first exposure (lag 0), the lag between the first and second exposure (lag 1), the lag between the second and third exposure (lag 2), and the lag between the third exposure and the final memory phase (lag 3). The first scan session of the continuous recognition phase for each participant corresponds to Day 0. Because memory is observed to abide by an exponential rule rather than linear time (Rubin & Wenzel, 1996), all temporal lags were quantified by expressing time intervals in seconds and transforming these intervals with the natural logarithm. Lag effects were then tested using mixed-effects regression models with either recognition confidence or temporal memory precision as a dependent variable and with each temporal lag as a separate predictor.

Representational similarity analyses

Representational similarity analyses were conducted on functional data (single-trial betas) from the continuous recognition phase, and were performed by assessing patterns of neural activity across voxels within each ROI evoked during single trials. Pattern similarity of all possible exposure pairings (Figure 3.3b; r(E1, E2), r(E2, E3), and r(E1, E3)) for each image was computed using Pearson correlation. The resulting correlation coefficients were then Fisher-transformed for further analyses. To avoid potential contamination of similarity from scanner-induced autocorrelation of signals, only correlations between image exposures that occurred across runs were considered (range of the trials excluded for each participant: 12-35).

Image-specificity analyses

We used two approaches to assess image-specificity in CA1 and entorhinal representations that predicted temporal memory.

Intact versus shuffled pattern similarity analysis.

Our first analysis tested whether temporal memory precision was predicted by image-specific pattern similarity (restricted to E1-E2 similarity) in CA1 and ERC using images tested in the temporal memory test (which were a subset of the full image set). Specifically, we randomly shuffled the E1-E2 mappings within each participant, such that each image's E1 was paired with a different image's E2. We then computed the pattern similarity of these shuffled exposure pairs and the new corresponding temporal lags. The shuffled E1-E2 pattern similarity scores and temporal lag information were then submitted to a mixed-effects logistic regression model predicting temporal memory precision. This procedure was performed 1,000 times, resulting in a null distribution of pattern similarity effects (betas values) for each ROI.

Target versus foil pattern similarity analysis.

Our second approach examined whether pattern similarity effects observed in CA1 and ERC were specific to individual images or were driven by general memory-related processes that could be shared across different images and/or differences in coarse temporal information (i.e., session effects). To do this, for each image included in the temporal memory test (a 'target'), we identified control images ('foils') according to two criteria: (1) targets and foils shared the same E1/E2 session number, but not run number, respectively (Figure 3.5a); (2) to control for generic memory states (recognition memory performance at each encounter), foils had to receive the same memory judgments as targets (i.e., to be responded correctly all three times), which were correctly

rejected at E1 and hit at E2 and E3. We then computed pattern similarity between target E1 and target E2 ('target similarity') and target E1 and foils E2 ('foil similarity'). This selection procedure resulted in different numbers of foils for each target image. For images with two or more foils, we used the median value of those foil similarity scores. To index the extent to which pattern similarity captures image-specific representations, foil similarity was subtracted from target similarity for each image (target similarity – foil similarity). This difference score between target and foil similarity was then submitted to a mixed-effects logistic regression model as a predictor of temporal memory precision, where a significant positive relationship would indicate that the pattern similarity that predicted temporal memory precision was driven by image-specific representations.

Statistical analyses

Behavioral and fMRI data were analyzed using a combination of permutation tests, paired *t* tests, repeated-measures ANOVA and mixed-effects regression models. Trial-level relationships between similarity measures and final memory performance were tested with mixed-effects linear/logistic regression models (for recognition confidence and temporal memory precision, respectively). For all permutation analyses, we used 1,000 permutations and assessed significance by computing the proportion of values in the null distribution that were higher/lower than the observed values. All *t* tests were two-tailed. For mixed-effects regression models, we used the participant as a random effect and other variables as fixed effects. A threshold of $p < 0.05$ was used to establish statistical significance for all analyses. fMRI analyses were corrected for multiple comparisons with Bonferroni corrections when applicable. Only ROIs that survived correction are reported except where otherwise noted.

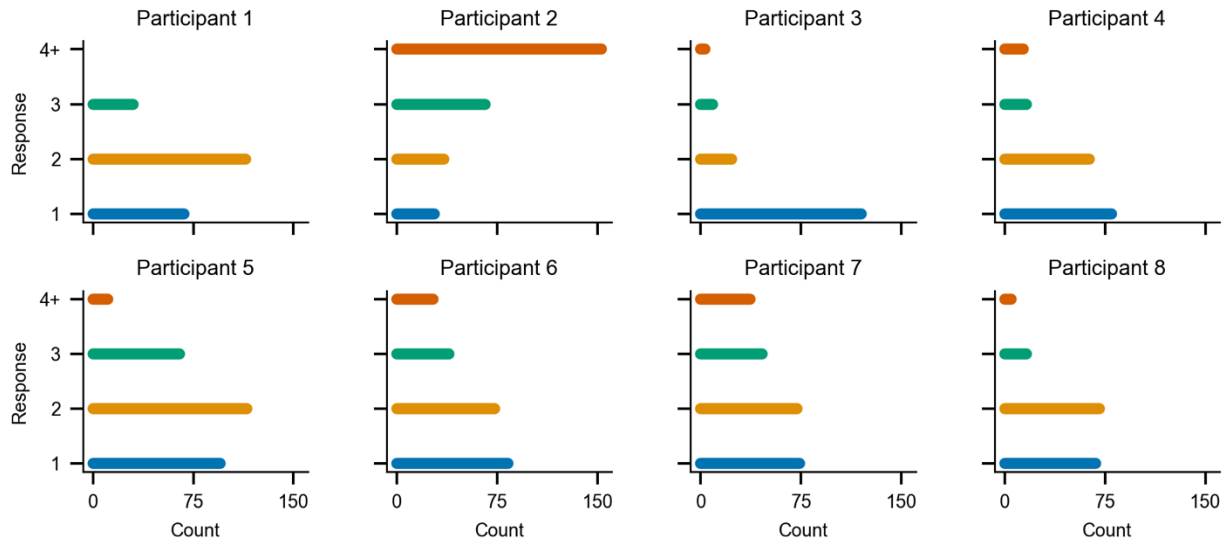

Figure 3.6. Subjective estimates of image frequency in the final memory phase. During the frequency test, participants were asked to indicate how many times they had seen each image (1, 2, 3, or 4 or more times). No participant preferentially chose the correct response (i.e., three times).

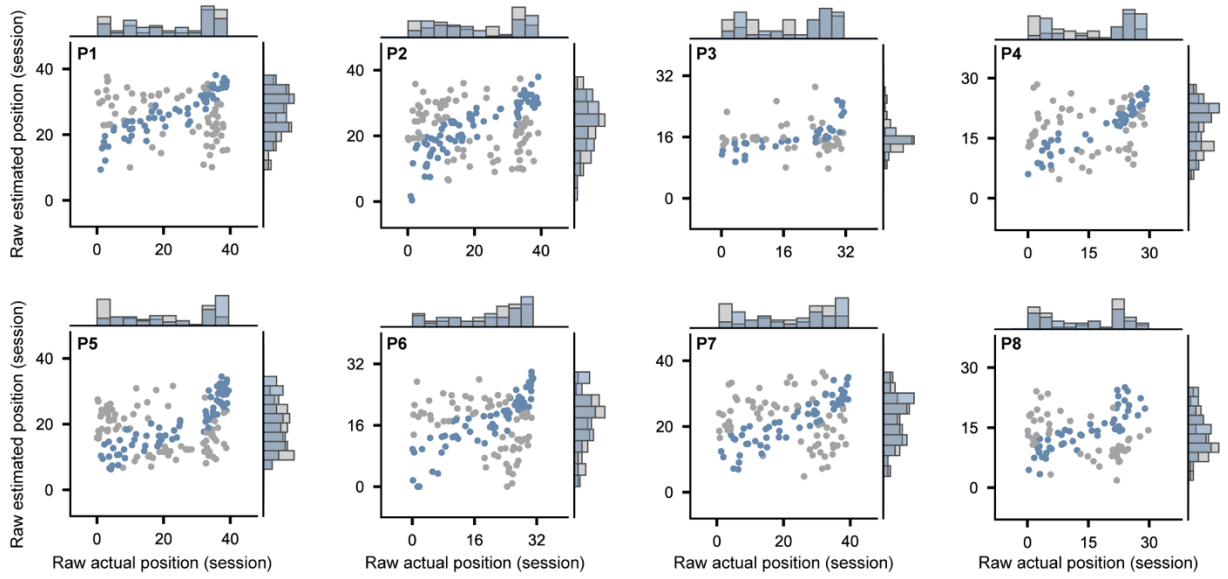

Figure 3.7. Raw estimated and actual temporal positions for each participant. Each dot indicates a temporal memory judgment. Temporal judgments for each participant were divided into 'high temporal precision' and 'low temporal precision' on the basis of temporal memory error (median split). Histograms on sides indicate distributions of raw estimated (vertical) and actual

(horizontal) temporal positions (overlaid across high- and low- temporal precision). Colors denote temporal memory precision (blue: high temporal precision; gray: low temporal precision).

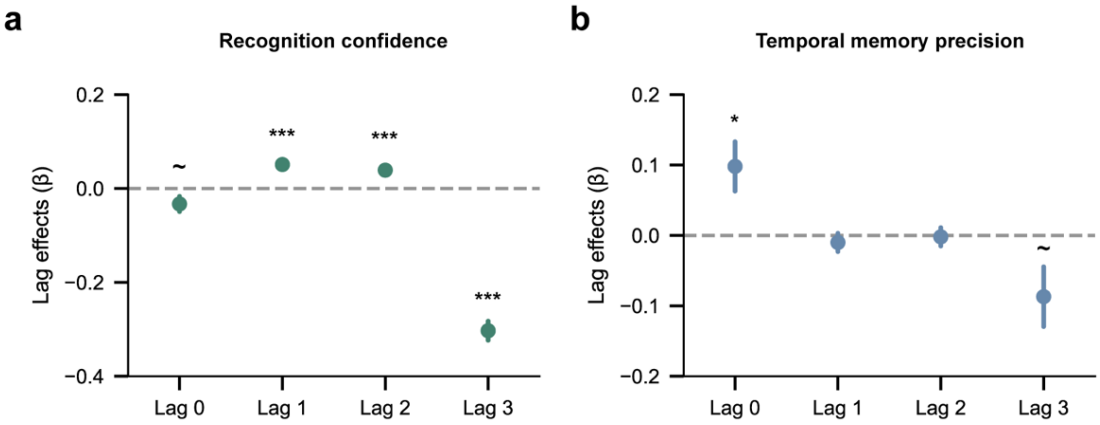

Figure 3.8. Lag effects on behavioral measures of final memory. (a) Lag effects on recognition confidence. There were spacing effects in recognition confidence such that confidence increased as a function of increasing temporal lags between exposures (lag 1: $\beta = 0.051$, $p < 0.001$; lag 2: $\beta = 0.039$, $p < 0.001$; in a mixed-effects linear regression model with recognition confidence as dependent variable and with each temporal lag as a separate predictor), and a forgetting effect wherein recognition confidence diminished as the lag between the last exposure and the final memory test increased (lag 3: $\beta = -0.303$, $p < 0.001$). **(b)** Lag effects on temporal memory precision. There was a recency effect across months with temporal memory precision increasing the later in the experiment an image was first encountered (lag 0: $\beta = 0.098$, $p = 0.021$; in a mixed-effects logistic regression model with temporal memory precision as dependent variable and with each temporal lag as a separate predictor). Error bars reflect mean ± s.e.m.; $\sim p < 0.10$; $*p < 0.05$; $***p < 0.001$.

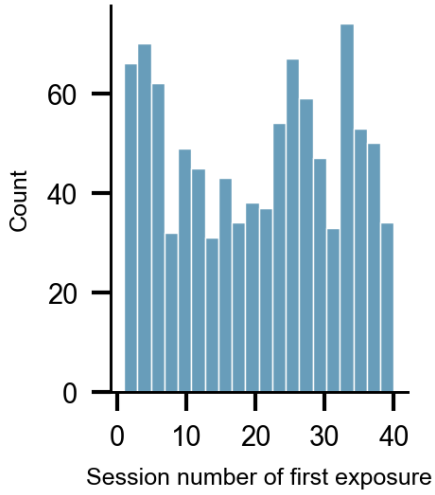

Figure 3.9. Distribution of images' first exposures. Counts of first exposures (E1) for old images included in analyses as a function of session number. Note: counts are summed across participants.

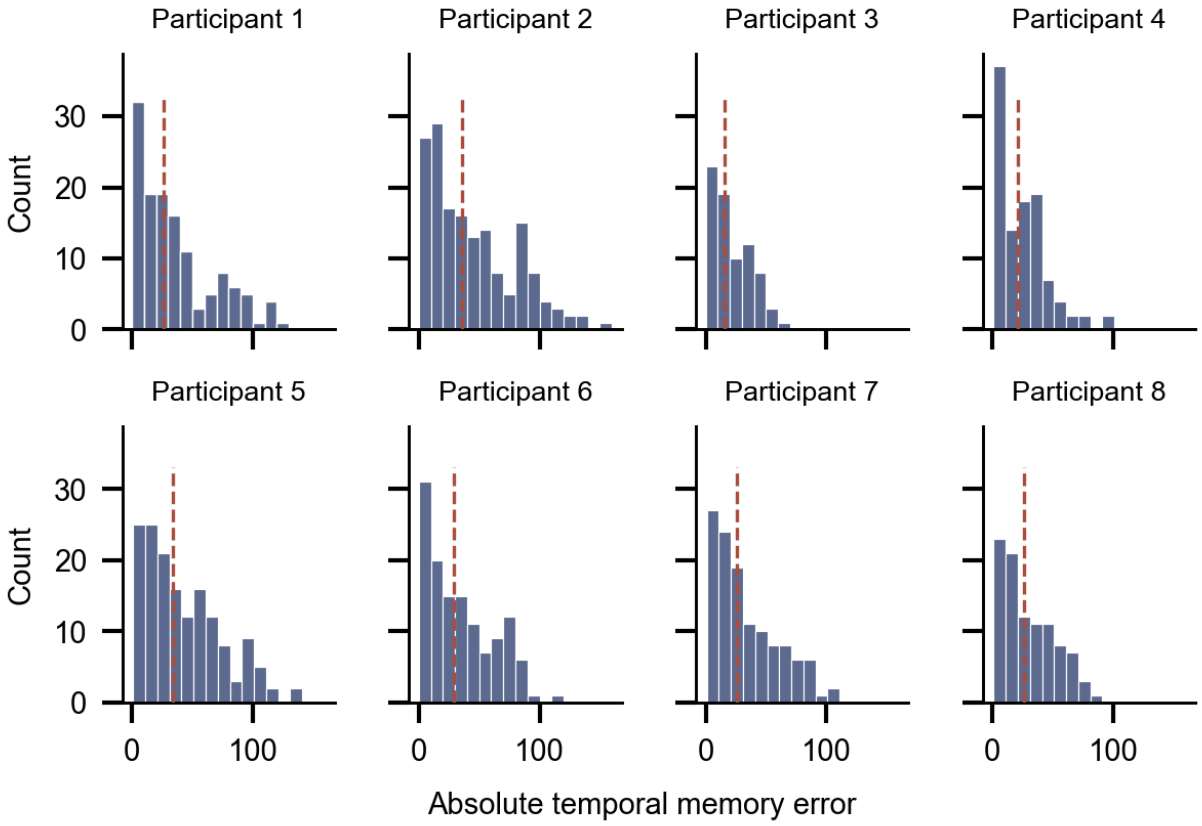

Figure 3.10. Distribution of individual participant's temporal memory error. Item-wise temporal memory error was calculated using the absolute difference between the ranked estimated temporal position and the ranked actual position. Red line denotes median value of the absolute temporal error. As the resulting distribution was highly non-normal, a median split was performed for all subsequent analyses of temporal error as a factor.

Table 3.1. Whole-brain representational similarity analysis in cortical ROIs. A mixed-effects logistic regression predicting temporal memory with pattern similarity as the main fixed effect of interest was conducted for each parcel of the HCP-MMP1 atlas to determine whether cortical regions outside of the MTL also exhibit pattern similarity effects on temporal memory. This table summarizes the whole-brain parcel level analysis with cortical ROIs showing significant relationships between pattern similarity and temporal memory before correcting for multiple comparisons ($p(uncorrected) < 0.05$). No regions survived correction for multiple comparisons.

Cortical ROIs (HCP-MMP1 atlas)	*z*-value	*p*-value (uncorrected)
L Frontal Eye Fields	2.027	0.043
L Supplementary and Cingulate Eye Field	2.693	0.007
L Area 6m anterior	2.761	0.006
L Area IFJp	2.356	0.018
L Area IFSp	2.043	0.041
L Area posterior 9-46v	2.020	0.043
L Entorhinal Cortex	2.001	0.045
L Perirhinal Ectorhinal Cortex	2.912	0.004
L Area 31pd	-2.011	0.044
R Area 31p ventral	-2.369	0.018
R Frontal Opercular Area 1	-2.062	0.039
R Entorhinal Cortex	2.082	0.037
R Area TG dorsal	2.109	0.035
R AreaTemporoParietoOcci pital Junction 1	2.119	0.034
R Area 31a	-1.963	0.050

References Cited

Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., Nau, M., Caron, B., Pestilli, F., Charest, I., Hutchinson, J. B., Naselaris, T., & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nature Neuroscience*, *25*(1), Article 1. https://doi.org/10.1038/s41593-021-00962-x

Bellmund, J. L. S., Deuker, L., Montijn, N. D., & Doeller, C. F. (2022). Mnemonic construction and representation of temporal structure in the hippocampal formation. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-30984-3

Bellmund, J. L. S., Polti, I., & Doeller, C. F. (2020). Sequence Memory in the Hippocampal–Entorhinal Region. *Journal of Cognitive Neuroscience*, *32*(11), 2056–2070. https://doi.org/10.1162/jocn_a_01592

Berron, D., Vieweg, P., Hochkeppler, A., Pluta, J. B., Ding, S.-L., Maass, A., Luther, A., Xie, L., Das, S. R., Wolk, D. A., Wolbers, T., Yushkevich, P. A., Düzel, E., & Wisse, L. E. M. (2017). A protocol for manual segmentation of medial temporal lobe subregions in 7Tesla MRI. *NeuroImage: Clinical*, *15*, 466–482. https://doi.org/10.1016/j.nicl.2017.05.022

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), 1–12. https://doi.org/10.1038/s41467-019-13377-x

Copara, M. S., Hassan, A. S., Kyle, C. T., Libby, L. A., Ranganath, C., & Ekstrom, A. D. (2014). Complementary Roles of Human Hippocampal Subregions during Retrieval of Spatiotemporal Context. *Journal of Neuroscience*, *34*(20), 6834–6842. https://doi.org/10.1523/JNEUROSCI.5341-13.2014

Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, *16*(6), 693–700. https://doi.org/10.1016/j.conb.2006.10.012

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99. https://doi.org/10.1016/j.tics.2014.12.004

Deuker, L., Bellmund, J. L., Navarro Schröder, T., & Doeller, C. F. (2016). An event map of memory space in the hippocampus. *eLife*, *5*, e16534. https://doi.org/10.7554/eLife.16534

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–386. https://doi.org/10.1016/j.tics.2007.08.001

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), Article 1. https://doi.org/10.1038/s41467-017-02752-1

DuBrow, S., & Davachi, L. (2014). Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation. *Journal of Neuroscience*, *34*(42), 13998–14005. https://doi.org/10.1523/JNEUROSCI.2535-14.2014

DuBrow, S., & Davachi, L. (2017). Commentary: Distinct neural mechanisms for remembering when an event occurred. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00189

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

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*(11), Article 11. https://doi.org/10.1038/nrn3827

Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., & Lipton, P. (2012). Towards a functional organization of episodic memory in the medial temporal lobe. *Neuroscience & Biobehavioral Reviews*, *36*(7), 1597–1608. https://doi.org/10.1016/j.neubiorev.2011.07.006

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The Medial Temporal Lobe and Recognition Memory. *Annual Review of Neuroscience*, *30*(1), 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Ezzyat, Y., & Davachi, L. (2014). Similarity breeds proximity: Pattern similarity within and across contexts is related to later mnemonic judgments of temporal proximity. *Neuron*, *81*(5), 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042

Ezzyat, Y., Kragel, J. E., Burke, J. F., Levy, D. F., Lyalenko, A., Wanda, P., O'Sullivan, L., Hurley, K. B., Busygin, S., Pedisich, I., Sperling, M. R., Worrell, G. A., Kucewicz, M. T., Davis, K. A., Lucas, T. H., Inman, C. S., Lega, B. C., Jobst, B. C., Sheth, S. A., … Kahana, M. J. (2017). Direct Brain Stimulation Modulates Encoding States and Memory Performance in Humans. *Current Biology*, *27*(9), 1251–1258. https://doi.org/10.1016/j.cub.2017.03.028

Folkerts, S., Rutishauser, U., & Howard, M. W. (2018). Human Episodic Memory Retrieval Is Accompanied by a Neural Contiguity Effect. *Journal of Neuroscience*, *38*(17), 4200–4211. https://doi.org/10.1523/JNEUROSCI.2312-17.2018

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, *113*(1), 44. https://doi.org/10.1037/0033-2909.113.1.44

Friedman, W. J. (2004). Time in Autobiographical Memory. *Social Cognition*, *22*(5), 591–605. https://doi.org/10.1521/soco.22.5.591.50766

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), Article 7615. https://doi.org/10.1038/nature18933

Hinrichs, J. V. (1970). A two-process memory-strength theory for judgment of recency. *Psychological Review*, *77*(3), 223–233. https://doi.org/10.1037/h0029101

Hintzman, D. L. (2005). Memory strength and recency judgments. *Psychonomic Bulletin & Review*, *12*(5), 858–864. https://doi.org/10.3758/BF03196777

Howard, M. W. (2017). Temporal and spatial context in the mind and brain. *Current Opinion in Behavioral Sciences*, *17*, 14–19. https://doi.org/10.1016/j.cobeha.2017.05.022

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General*, *142*(4), 1211. https://doi.org/10.1037/a0033621

Howard, M. W., & Kahana, M. J. (2002). A Distributed Representation of Temporal Context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Howard, M. W., Shankar, K. H., Aue, W. R., & Criss, A. H. (2015). A distributed representation of internal time. *Psychological Review*, *122*(1), 24–53. https://doi.org/10.1037/a0037840

Howard, M. W., Viskontas, I. V., Shankar, K. H., & Fried, I. (2012). Ensembles of human MTL neurons "jump back in time" in response to a repeated stimulus. *Hippocampus*, *22*(9), 1833–1847. https://doi.org/10.1002/hipo.22018

Hsieh, L.-T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal Activity Patterns Carry Information about Objects in Temporal Context. *Neuron*, *81*(5), 1165–1178. https://doi.org/10.1016/j.neuron.2014.01.015

Jenkins, L. J., & Ranganath, C. (2010). Prefrontal and Medial Temporal Lobe Activity at Encoding Predicts Temporal Context Memory. *Journal of Neuroscience*, *30*(46), 15558–15565. https://doi.org/10.1523/JNEUROSCI.1337-10.2010

Jenkins, L. J., & Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. *Hippocampus*, *26*(5), 554–559. https://doi.org/10.1002/hipo.22571

Kragel, J. E., Ezzyat, Y., Lega, B. C., Sperling, M. R., Worrell, G. A., Gross, R. E., Jobst, B. C., Sheth, S. A., Zaghloul, K. A., Stein, J. M., & Kahana, M. J. (2021). Distinct cortical systems reinstate the content and context of episodic memories. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-24393-1

Kraus, B. J., Robinson, R. J., White, J. A., Eichenbaum, H., & Hasselmo, M. E. (2013). Hippocampal "Time Cells": Time versus Path Integration. *Neuron*, *78*(6), 1090–1101. https://doi.org/10.1016/j.neuron.2013.04.015

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), Article 4. https://doi.org/10.1038/nn.2498

Lee, A. C. H., Thavabalasingam, S., Alushaj, D., Çavdaroğlu, B., & Ito, R. (2020). The hippocampus contributes to temporal duration memory in the context of event sequences: A cross-species perspective. *Neuropsychologia*, *137*, 107300. https://doi.org/10.1016/j.neuropsychologia.2019.107300

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740–755). Springer.

Lositsky, O., Chen, J., Toker, D., Honey, C. J., Shvartsman, M., Poppenk, J. L., Hasson, U., & Norman, K. A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *eLife*, *5*, e16070. https://doi.org/10.7554/eLife.16070

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. *Neuron*, *71*(4), 737–749. https://doi.org/10.1016/j.neuron.2011.07.012

Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., & Leutgeb, J. K. (2012). Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, *109*(47), 19462–19467. https://doi.org/10.1073/pnas.1214107109

Manning, J. R., Polyn, S. M., Baltuch, G. H., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences*, *108*(31), 12893–12897. https://doi.org/10.1073/pnas.1015174108

Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron*, *56*(3), 530–540. https://doi.org/10.1016/j.neuron.2007.08.017

Mau, W., Sullivan, D. W., Kinsky, N. R., Hasselmo, M. E., Howard, M. W., & Eichenbaum, H. (2018). The Same Hippocampal CA1 Population Simultaneously Codes Temporal Information over Multiple Timescales. *Current Biology*, *28*(10), 1499-1508.e4. https://doi.org/10.1016/j.cub.2018.03.051

Montchal, M. E., Reagh, Z. M., & Yassa, M. A. (2019). Precise temporal memories are supported by the lateral entorhinal cortex in humans. *Nature Neuroscience*, *22*(2), 284–288. https://doi.org/10.1038/s41593-018-0303-1

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

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsáki, G. (2008). Internally Generated Cell Assembly Sequences in the Rat Hippocampus. *Science*, *321*(5894), 1322–1327. https://doi.org/10.1126/science.1159775

Polyn, S. M., & Cutler, R. A. (2017). Retrieved-context models of memory search and the neural representation of time. *Current Opinion in Behavioral Sciences*, *17*, 203–210. https://doi.org/10.1016/j.cobeha.2017.09.007

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129–156. https://doi.org/10.1037/a0014420

Ranganath, C. (2010). Binding Items and Contexts: The Cognitive Neuroscience of Episodic Memory. *Current Directions in Psychological Science*, *19*(3), 131–137. https://doi.org/10.1177/0963721410368805

Ranganath, C., & Hsieh, L.-T. (2016). The hippocampus: A special place for time. *Annals of the New York Academy of Sciences*, *1369*(1), 93–110. https://doi.org/10.1111/nyas.13043

Reddy, L., Zoefel, B., Possel, J. K., Peters, J., Dijksterhuis, D. E., Poncet, M., Straaten, E. C. W. van, Baayen, J. C., Idema, S., & Self, M. W. (2021). Human Hippocampal Neurons Track Moments in a Sequence of Events. *Journal of Neuroscience*, *41*(31), 6714–6725. https://doi.org/10.1523/JNEUROSCI.3157-20.2021

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

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, *103*(4), 734–760. https://doi.org/10.1037/0033-295X.103.4.734

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of Object Representations in the Human Medial Temporal Lobe Based on Temporal Regularities. *Current Biology*, *22*(17), 1622–1627. https://doi.org/10.1016/j.cub.2012.06.056

Sederberg, P. B., Gershman, S. J., Polyn, S. M., & Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin & Review*, *18*(3), 455–468. https://doi.org/10.3758/s13423-011-0086-9

Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The Medial Temporal Lobe. *Annual Review of Neuroscience*, *27*(1), 279–306. https://doi.org/10.1146/annurev.neuro.27.070203.144130

Tambini, A., & Davachi, L. (2019). Awake Reactivation of Prior Experiences Consolidates Memories and Biases Cognition. *Trends in Cognitive Sciences*, *23*(10), 876–890. https://doi.org/10.1016/j.tics.2019.07.008

Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M.-B., & Moser, E. I. (2018). Integrating time from experience in the lateral entorhinal cortex. *Nature*, *561*(7721), 57–62. https://doi.org/10.1038/s41586-018-0459-6

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, *53*, 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B. E., Sperling, M., & Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proceedings of the National Academy of Sciences*, *117*(45), 28463–28474. https://doi.org/10.1073/pnas.2013250117

Wang, F., & Diana, R. A. (2017). Temporal context in human fMRI. *Current Opinion in Behavioral Sciences*, *17*, 57–64. https://doi.org/10.1016/j.cobeha.2017.06.004

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269. https://doi.org/10.1146/annurev.psych.55.090902.141555

Yaffe, R. B., Kerr, M. S. D., Damera, S., Sarma, S. V., Inati, S. K., & Zaghloul, K. A. (2014). Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal dynamics during successful memory retrieval. *Proceedings of the National Academy of Sciences*, *111*(52), 18727–18732. https://doi.org/10.1073/pnas.1417017112

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

CHAPTER IV

RECONSTRUCTING THE TEMPORAL CONTEXT OF PAST EVENTS

This chapter is work in progress by Futing Zou, J. Benjamin Hutchinson, and Brice A. Kuhl.

Abstract

Leading theories of memory propose that our experiences are embedded within slowly drifting representations that capture the passage of time (temporal context). When events from the past are remembered, temporal context representations are thought to also be reinstated. Here, using natural language processing methods and inverted fMRI encoding models, we developed a novel approach to directly measure the reinstatement of temporal context. Specifically, we show that when a previously-encountered stimulus is re-encountered, activity patterns in ventromedial prefrontal cortex reflect the semantic information of temporally-adjacent stimuli of its original encounter. That is, re-encountering a stimulus reinstates semantic information that were encoded close in time when the stimulus was originally encountered. Moreover, we show that such reinstatement depends on the degree to which CA1 activity patterns from the first encounter with a stimulus were re-expressed during re-encounters. Together, these findings constitute novel evidence of temporal context reinstatement and highlight the influence of past events on ongoing processing.

Introduction

Central to leading theories and computational models of episodic memory is the idea that individual events are embedded in a broader "temporal context" (Howard & Kahana, 2002; Norman et al., 2001). When a past event is retrieved from memory, this is thought to trigger

reinstatement of the event's temporal context (Polyn et al., 2009; Sederberg et al., 2008). As such, remembering an event not only recovers the event itself, but also reinstates information related to other events that were encoded nearby in time. However, methods for directly measuring temporal context reinstatement are currently limited.

One way in which temporal context reinstatement has been measured is through reinstatement of neural activity patterns. This approach is motivated by the idea that temporal context is encoded within distributed, drifting patterns of neural activity. Indeed, converging evidence from human neuroimaging studies has shown that recalling a stimulus not only involves reinstatement of neural activity patterns that were initially engaged during the encoding of that stimulus, but also reinstatement of the activity patterns that were temporally-adjacent to encoding of that stimulus (Folkerts et al., 2018; Manning et al., 2011; Yaffe et al., 2014). While important, these neural activity patterns, on their own, do not indicate the nature of the information that is reinstated.

Human and rodent studies have consistently implicated hippocampal subfield CA1 as being particularly important for encoding and remembering temporal information (Davachi & DuBrow, 2015; Eichenbaum, 2014; Ranganath & Hsieh, 2016; Tsao et al., 2022). Importantly, in a recent human fMRI study (Chapter III) (Zou et al., 2023), we found that, across stimulus encounters, similarity of distributed, event-specific activity patterns in CA1 positively predicted subsequent temporal memory for the original encounter, suggesting that when a stimulus is re-encountered in a new temporal context, the stimulus' original temporal context is reinstated, thereby benefiting subsequent temporal memory. While this study strongly supports the idea of temporal context reinstatement (in CA1), neural pattern similarity is not, as noted above, an unambiguous measure

of reinstatement. Thus, this interpretation would benefit from additional evidence—particularly using methods that more directly measure temporal context reinstatement.

Here rather than measuring reinstatement of neural activity patterns, we developed and used a novel fMRI-based approach to directly test for reinstatement of information that were temporally-adjacent to initial encoding of a stimulus. Specifically, we used natural language processing methods to characterize the content of natural scene images and inverted fMRI encoding models to test whether repetition of a stimulus reinstated the semantic content that preceded and followed the stimulus' original encounter. Thus, temporal context reinstatement was operationalized as the degree to which neural patterns evoked by a stimulus' repetition contain information about the semantic content that were encoded close in time when the stimulus was originally encountered. Moreover, motivated by our recent study (Zou et al., 2023) and by evidence for temporal processing in CA1 (MacDonald et al., 2011; Mankin et al., 2012; Mau et al., 2018), we tested whether similarity of CA1 activity patterns across stimulus repetitions related to the degree of temporal context reinstatement. This approach allowed us to explicitly test predictions from leading computational models and theories of episodic memory and hippocampal function.

Methods

Natural Scenes Dataset

To test our hypothesis, we analyzed data from the massive Natural Scenes Dataset (Allen et al., 2022). Eight participants performed a continuous recognition task in which they viewed thousands of natural scene images that were repeatedly presented across many scan sessions, from which the first two encounters of each stimulus (E1, E2) were used for the present study. Analyses were restricted to E1 and E2 trials that were drawn from the same session (but different scan runs)

and were each associated with correct behavioral responses (i.e., E1 = 'new' responses, E2 = 'old' responses) to avoid potential confounds. Images used in NSD were taken from the Microsoft Common Objects in Context (COCO) database (Lin et al., 2014), which includes human-generated sentence captions describing each image.

Region of interest (ROI) definition

To probe whether neural activity patterns evoked by a stimulus' repetition contain information about the semantic content that were temporally-adjacent to initial encoding of that stimulus, we focused on three regions of interest (ROIs) that we predicted would be sensitive to the semantic content within scene images: bilateral ventromedial prefrontal cortex (vmPFC), bilateral angular gyrus (AG), and bilateral lateral occipitotemporal cortex (LOTC). All ROIs were drawn from the surface-based Human Connectome Project multimodal parcellation (HCP-MMP) atlas of human cortical areas (Glasser et al., 2016). To probe the relationship between neural pattern similarity in CA1 and temporal context reinstatement, we used bilateral CA1 that were manually drawn on the high-resolution T2 images obtained for each participant.

Inverted encoding model

We operationalized the temporal context of E1 as the semantic information of the stimulus that temporally-adjacent to E1 (i.e., E1-1, E1+1). The primary goal of our analyses was to reconstruct the semantic content of stimuli that neighbored E1 based on activity patterns evoked during E2. To characterize the semantic content of scene images, we transformed independent annotations (text descriptions) of the scene images into 512-dimensional semantic embeddings using Google's Universal Sentence Encoder (USE) (Cer et al., 2018). For each stimulus, we

operationalized E1's temporal context as the mean of semantic embeddings for the stimuli that neighbored E1 (i.e., E1-1, E1+1). We then applied principal component analysis (PCA) on the resulting semantic embeddings and used the top 20 PCs as a 20-dimensional representation of the semantic content of each scene image. These PC scores were used for the inverted encoding model described below.

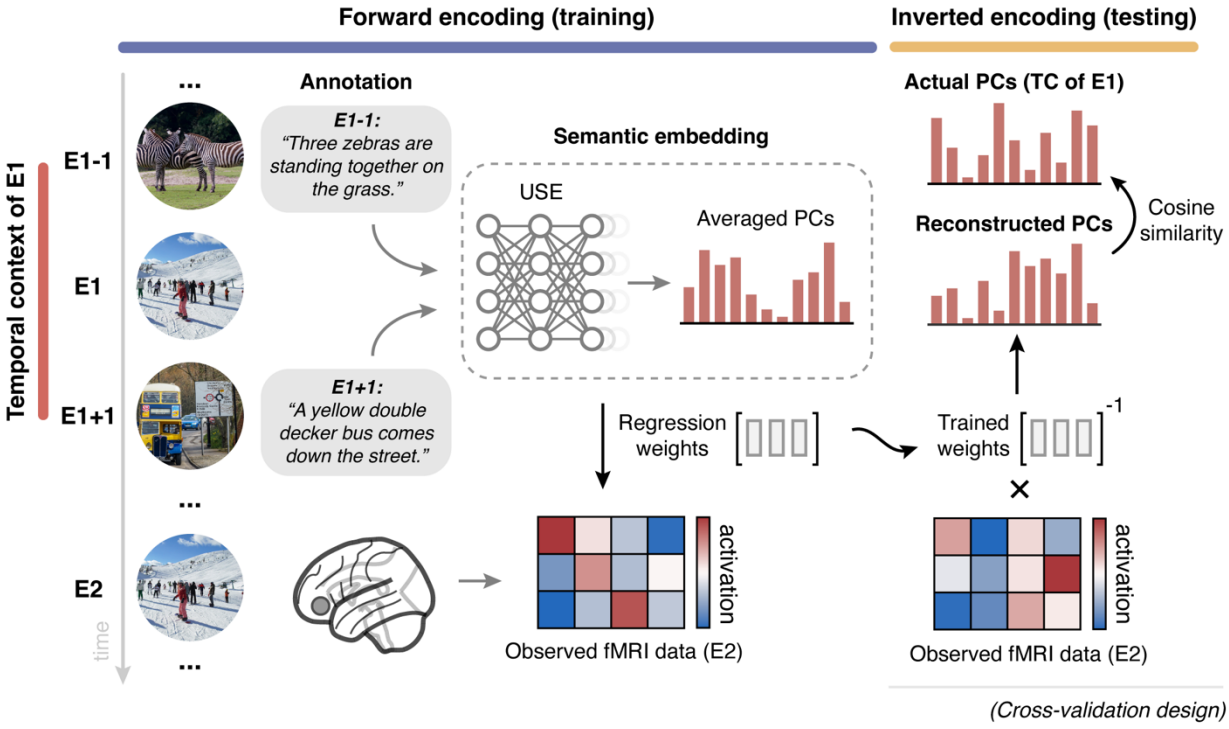

Figure 4.1. Data analysis. A schematic of the inverted encoding models used for reconstructing semantic components of scenes from fMRI activity patterns. To characterize the semantic content of scene images, independent annotations (text descriptions) of the scene images were transformed into 512-dimensional semantic embeddings using Google's Universal Sentence Encoder (USE) (Cer et al., 2018). For each image, the top 20 principle components (PCs) of the mean of semantic embeddings for the stimuli that neighbored E1 (i.e., E1-1, E1+1) was used as the proxy for E1's temporal context. A linear mapping from the semantic components (of E1-1 and E1+1) to fMRI activity patterns evoked during E2 was obtained using ridge regression. This encoding model was inverted and applied to held-out data (leave-one-session-out) in a cross-validated manner to reconstruct semantic components. Reconstruction accuracy was computed as the cosine similarity between the reconstructed and actual semantic components.

Reconstructions of semantic components were generated using a cross-validation approach (Figure 4.1). Specifically, using ridge regression, we first learned a direct linear mapping from the semantic components (of E1-1 and E1+1) to fMRI activity patterns evoked during E2. We then inverted this encoding model and applied it to held-out data (leave-one-session-out) in a cross-validated manner to reconstruct semantic components. Reconstruction accuracy was computed as the cosine similarity between the reconstructed and actual semantic components. Above-chance reconstruction of semantic embeddings neighboring E1 would indicate evidence that E2 triggers reinstatement of E1's temporal context.

Results

Here we focused our analyses on three regions of interest (ROIs) that we predicted would be sensitive to the semantic content within scene images: (1) ventromedial prefrontal cortex (vmPFC) which forms high-level schemas (Gilboa & Marlatte, 2017), (2) angular gyrus (AG) which forms event-specific memory representations (Humphreys et al., 2021), and (3) lateral occipitotemporal cortex (LOTC) which is sensitive to current visual content (Konkle & Caramazza, 2013). We first tested whether activity patterns at E2 contained information about the current stimulus (E2). Consistent with previous studies (Lee & Kuhl, 2016; Wang et al., 2023), robust content reconstruction was obtained from all ROIs ($ps < 0.001$; permutation test).

Of critical interest was whether semantic information that neighbored E1 (i.e., the average semantic embeddings for E1-1 and E1+1) could be reconstructed from activity patterns elicited at E2. Successful reconstruction would indicate that participants reinstated the temporal context of E1 when E2 was encountered. Strikingly, we found that average semantic embeddings of E1-1 and E1+1 were successfully reconstructed from E2 activity patterns in vmPFC (Figure 4.2; $p = 0.035$;

permutation test), with above-chance reconstruction for 3/8 individual participants ($p = 0.006$; binomial test). While AG and LOTC showed sensitivity to the content of the current stimulus (E2), neither region supported reconstruction of E1-1 (Figure 4.2; $ps > 0.19$). Together, these results demonstrate that, when a stimulus was re-encountered, vmPFC supported successful reconstruction of semantic content from stimuli that were temporally-adjacent to its original encounter, constituting novel evidence for temporal context reinstatement.

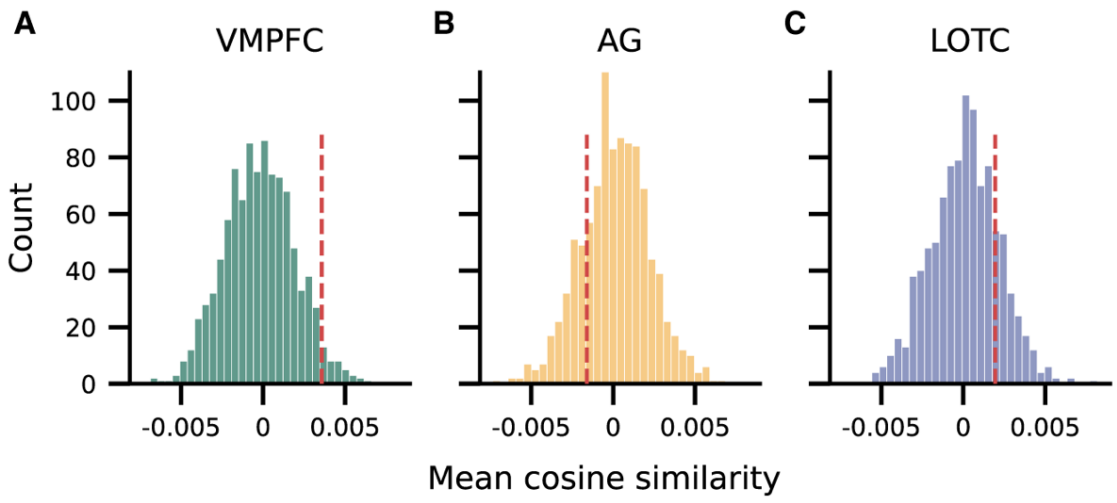

Figure 4.2. Content reconstruction accuracy. Mean cosine similarity between reconstructed and actual semantic components for each ROI. Red dashed lines depict actual similarity values.

Motivated by our recent work (Zou et al., 2023), we next sought to test whether neural pattern similarity in CA1 across stimulus encounters (E1-E2 similarity) relates to the degree of temporal context reinstatement (i.e., content reconstruction (in vmPFC) of semantic information that neighbored E1 during E2). In other words, do CA1 representations of repeated events reflect the extent that temporal context is reinstated? Specifically, to ensure that pattern similarity did not reflect generic cognitive processes, we used a measure of stimulus-specific similarity in which we compared 'within-image' pattern similarity to 'across-image' pattern similarity for each image. We

then compared reconstruction accuracy in vmPFC for images associated with high versus low stimulus-specific similarity in CA1 (based on a participant-specific median split). Statistical significance of the reconstruction accuracy difference between high and low CA1 similarity was evaluated using a permutation test that shuffles the images' CA1 similarity identities within each participant. Strikingly, we found that reconstruction accuracy in vmPFC was significantly greater for images with high CA1 similarity relative to low CA1 similarity (Figure 4.3; $p = 0.038$; permutation tests). Specifically, reconstruction accuracy in vmPFC did not differ from zero for images with low CA1 similarity (Figure 4.3; $t = 0.10$, $p = 0.46$, one-sample t-test), but was significantly greater than zero for > 24 h spacing (Figure 4.3; $t = 2.44$, $p = 0.022$, one-sample t-test). Thus, successful reconstruction in vmPFC only emerged when CA1 similarity was relatively high. Together, the results provide evidence that temporal context reinstatement is predicted by the degree to which CA1 activity patterns from the first encounter with an image were re-expressed during re-encounters.

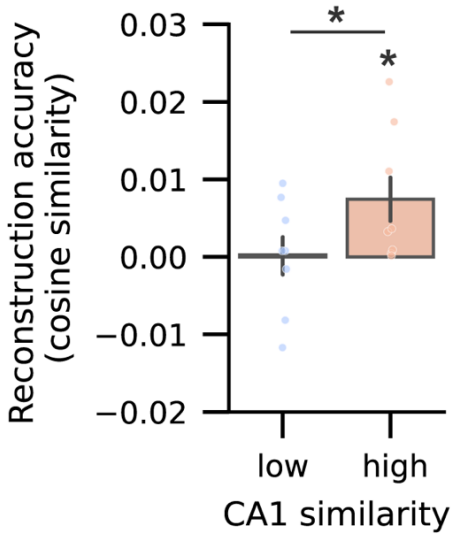

Figure 4.3. Reconstruction accuracy difference between high- and low-similarity in CA1 (median split). Reconstruction accuracy in vmPFC was significantly greater for images with high CA1 similarity relative to low CA1 similarity (low versus high: $p = 0.038$; permutation tests, $n = 1{,}000$). Reconstruction accuracy did not differ from zero for images with low CA1 similarity ($t = 0.10$, $p = 0.46$, one-sample t-test) but was significantly greater than zero for images with high CA1 similarity ($t = 2.44$, $p = 0.022$, one-sample t-test). Dots denote independent participants ($n = 8$); $^*p < 0.05$.

Discussion

In this study, we employed inverted semantic encoding models to directly measure temporal context reinstatement and tested how temporal context reinstatement relate to representations in CA1. Our results are two-fold. First, we found that, when a stimulus was re-encountered, vmPFC supported successful reconstruction of information that were encoded nearby in time when the stimulus was first encountered. This represents novel evidence for temporal context reinstatement, specifically demonstrating reinstatement of semantic content from temporally-adjacent stimuli. Second, we found that temporal context reinstatement was associated with greater pattern similarity across stimulus encounters in CA1, suggesting that distributed patterns of activity in CA1 codes for and reinstates temporal context information. Together, our results provide fundamental new insight into the nature of temporal context reinstatement and highlight the influence of past events on ongoing processing.

Highly-influential models of episodic memory have operationalized temporal context as a blend of stimuli that neighbor in time as new stimuli are encoded (Howard & Kahana, 2002; Norman et al., 2001). In other words, stimuli are encoded in the context of neighboring stimuli. When the stimulus is re-encountered, it may trigger reinstatement of the stimuli that were encoded nearby in time. While this account is strongly supported by behavioral studies showing that stimuli tend to be recalled from memory alongside stimuli that were studied close together in time (Sederberg et al., 2008), actually measuring this in the brain is a challenge. Here, we developed a

novel measure of temporal context reinstatement by testing whether semantic content of stimuli neighboring its original encounter (E1) could be reconstructed based on patterns of fMRI activity evoked by subsequent encounter (E2). Indeed, we found evidence for successful reconstruction of semantic information from stimuli neighboring E1 from activity during E2, indicating that temporal context of E1 was reinstated when E2 was encountered. In contrast, whereas AG and LOTC exhibited sensitivity to the content of current images, neither region supported reconstruction of neighboring images from prior context. Thus, it raises the intriguing idea that temporal context reinstatement may preferentially lead to retrieval of high-level schemas from temporally-adjacent stimuli. Thus, in future work it will be of interest to further characterize the selectivity of the information that is reinstated as understanding it will help refine theories of temporal context and episodic memory.

Moreover, by relating patterns of activity across repeated stimulus encounters to measure of temporal context, we found that greater CA1 similarity was associated with better reconstruction accuracy. This finding is striking in light of accumulating evidence documenting time sensitive cells within rodent CA1 (Eichenbaum, 2014; Tsao et al., 2022) and implicating CA1 representations in preserving temporal memory (Zou et al., 2023). In particular, this result complements our previous finding that, when stimuli are re-encountered, memory for the original temporal context will be preserved to the extent that activity patterns in CA1 are similar to (or reinstate) the activity patterns expressed when the stimulus was first encountered (Zou et al., 2023) (also see Chapter III). Together with current evidence that CA1 representations of repeated events track the degree of temporal context reinstatement, we were able to provide new and comprehensive insight into the mechanisms through which CA1 contributes to the temporal organization of events in memory.

References Cited

Allen, E. J., St-Yves, G., Wu, Y., Breedlove, J. L., Prince, J. S., Dowdle, L. T., Nau, M., Caron, B., Pestilli, F., Charest, I., Hutchinson, J. B., Naselaris, T., & Kay, K. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience and artificial intelligence. *Nature Neuroscience*, *25*(1), Article 1. https://doi.org/10.1038/s41593-021-00962-x

Cer, D., Yang, Y., Kong, S., Hua, N., Limtiaco, N., John, R. S., Constant, N., Guajardo-Cespedes, M., Yuan, S., Tar, C., Sung, Y.-H., Strope, B., & Kurzweil, R. (2018). *Universal Sentence Encoder* (arXiv:1803.11175). arXiv. https://doi.org/10.48550/arXiv.1803.11175

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99. https://doi.org/10.1016/j.tics.2014.12.004

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*(11), Article 11. https://doi.org/10.1038/nrn3827

Folkerts, S., Rutishauser, U., & Howard, M. W. (2018). Human Episodic Memory Retrieval Is Accompanied by a Neural Contiguity Effect. *Journal of Neuroscience*, *38*(17), 4200–4211. https://doi.org/10.1523/JNEUROSCI.2312-17.2018

Gilboa, A., & Marlatte, H. (2017). Neurobiology of Schemas and Schema-Mediated Memory. *Trends in Cognitive Sciences*, *21*(8), 618–631. https://doi.org/10.1016/j.tics.2017.04.013

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), Article 7615. https://doi.org/10.1038/nature18933

Howard, M. W., & Kahana, M. J. (2002). A Distributed Representation of Temporal Context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Humphreys, G. F., Ralph, M. A. L., & Simons, J. S. (2021). A Unifying Account of Angular Gyrus Contributions to Episodic and Semantic Cognition. *Trends in Neurosciences*, *44*(6), 452–463. https://doi.org/10.1016/j.tins.2021.01.006

Konkle, T., & Caramazza, A. (2013). Tripartite Organization of the Ventral Stream by Animacy and Object Size. *Journal of Neuroscience*, *33*(25), 10235–10242. https://doi.org/10.1523/JNEUROSCI.0983-13.2013

Lee, H., & Kuhl, B. A. (2016). Reconstructing Perceived and Retrieved Faces from Activity Patterns in Lateral Parietal Cortex. *Journal of Neuroscience*, *36*(22), 6069–6082. https://doi.org/10.1523/JNEUROSCI.4286-15.2016

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740–755). Springer.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. *Neuron*, *71*(4), 737–749. https://doi.org/10.1016/j.neuron.2011.07.012

Mankin, E. A., Sparks, F. T., Slayyeh, B., Sutherland, R. J., Leutgeb, S., & Leutgeb, J. K. (2012). Neuronal code for extended time in the hippocampus. *Proceedings of the National Academy of Sciences*, *109*(47), 19462–19467. https://doi.org/10.1073/pnas.1214107109

Manning, J. R., Polyn, S. M., Baltuch, G. H., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences*, *108*(31), 12893–12897. https://doi.org/10.1073/pnas.1015174108

Mau, W., Sullivan, D. W., Kinsky, N. R., Hasselmo, M. E., Howard, M. W., & Eichenbaum, H. (2018). The Same Hippocampal CA1 Population Simultaneously Codes Temporal Information over Multiple Timescales. *Current Biology*, *28*(10), 1499-1508.e4. https://doi.org/10.1016/j.cub.2018.03.051

Norman, K. A., Detre, G., & Polyn, S. M. (2001). Computational Models of Episodic Memory. In R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (1st ed., pp. 189–225). Cambridge University Press. https://doi.org/10.1017/CBO9780511816772.011

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129–156. https://doi.org/10.1037/a0014420

Ranganath, C., & Hsieh, L.-T. (2016). The hippocampus: A special place for time. *Annals of the New York Academy of Sciences*, *1369*(1), 93–110. https://doi.org/10.1111/nyas.13043

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*(4), 893–912. https://doi.org/10.1037/a0013396

Tsao, A., Yousefzadeh, S. A., Meck, W. H., Moser, M.-B., & Moser, E. I. (2022). The neural bases for timing of durations. *Nature Reviews Neuroscience*, *23*(11), Article 11. https://doi.org/10.1038/s41583-022-00623-3

Wang, Y., Lee, H., & Kuhl, B. A. (2023). Mapping multidimensional content representations to neural and behavioral expressions of episodic memory. *NeuroImage*, *277*, 120222. https://doi.org/10.1016/j.neuroimage.2023.120222

Yaffe, R. B., Kerr, M. S. D., Damera, S., Sarma, S. V., Inati, S. K., & Zaghloul, K. A. (2014). Reinstatement of distributed cortical oscillations occurs with precise spatiotemporal

dynamics during successful memory retrieval. *Proceedings of the National Academy of Sciences*, *111*(52), 18727–18732. https://doi.org/10.1073/pnas.1417017112

Zou, F., Wanjia, G., Allen, E. J., Wu, Y., Charest, I., Naselaris, T., Kay, K., Kuhl, B. A., Hutchinson, J. B., & DuBrow, S. (2023). Re-expression of CA1 and entorhinal activity patterns preserves temporal context memory at long timescales. *Nature Communications*, *14*(1), 4350. https://doi.org/10.1038/s41467-023-40100-8

CHAPTER V

GENERAL DISCUSSION

Much of our everyday life is composed of repeated experiences. Repeated experiences not only allow opportunities for retrieving prior experiences that add up to improve later memory, but also entail the risk of creating competition between memories that can lead to interference. In this dissertation, I have characterized some of the neocortical and hippocampal mechanisms that preserve memories when experiences are repeated.

In Chapter II, we tested how spacing influences neural representations of repeated events that benefits memory. We found that spaced learning increased the similarity of distributed, stimulus-specific activity patterns in vmPFC. By systematically considering spacing from seconds to months, we found that these increases in vmPFC similarity not only parallelled, but predicted the behavioral benefits of spacing over long timescales. Finally, we found that the benefits of spaced learning—in brain and behavior—critically depended on remembering and re-encoding past experience. Taken together, we argue that spacing effect can be explained by a balance between retrieval and encoding processes.

In Chapter III, we tested whether neural pattern similarity across repeated stimulus encounters predicted subsequent temporal memory—specifically memory for when the stimulus was first encountered. In this study, the lags between stimulus repetitions ranged from seconds to many months and temporal memory were only made at the very end of the experiment (after 30-40 experimental sessions distributed across ~10 months). We found that greater neural pattern similarity across stimulus repetitions was associated with better (more precise) temporal memory for the original encounter. Notably, we observed these pattern similarity effects within two MTL subregions that have specifically been implicated in temporal processing and temporal memory:

hippocampal subfield CA1 (Davachi & DuBrow, 2015; Eichenbaum, 2014; Huerta et al., 2000; MacDonald et al., 2011; Mau et al., 2018) and entorhinal cortex (Goyal et al., 2018; Montchal et al., 2019; Tsao et al., 2018). These effects observed within regions implicated in temporal memory is consistent with the idea that when encoding a stimulus in a new temporal context, the stimulus' original temporal context is reinstated, thereby benefiting subsequent temporal memory. However, this interpretation would benefit from additional evidence—particularly using methods that more directly measure temporal context reinstatement.

In Chapter IV, we developed a novel approach for directly measuring temporal context reinstatement triggered by stimulus repetitions and characterizing what information gets reinstated. Specifically, using NLP methods and semantic encoding models, we found that, when a previously-encountered stimulus is re-encountered, vmPFC supported successful reconstructions of the semantic information of temporally-adjacent stimuli of its original encounter. That is, re-encountering a stimulus reinstates semantic information that were encoded close in time when the stimulus was originally encountered, a new marker of temporal context reinstatement. Critically, we found that such reinstatement depended on the degree to which CA1 activity patterns from the first encounter with a stimulus were re-expressed during re-encounters. Collectively, in Chapter III and IV, we advance an argument in which stimulus repetitions induce reinstatement of prior temporal contexts in CA1, thereby preserving precise temporal memory.

Thus, temporal context reinstatement is considered a mechanism that protects temporal memories when stimuli repeat. This is consistent with evidence from studies using classic memory interference paradigms in which a single stimulus is paired with multiple associates (A-B, A-C designs where A represents a repeated stimulus and B and C represent distinct associates). Specifically, using pattern-based fMRI analyses, several studies have shown that A-C encoding

trials can elicit spontaneous reinstatement of the B term (i.e., the associate previously encoded with A) (Chanales et al., 2019; Koen & Rugg, 2016; Kuhl et al., 2010; Richter et al., 2016; Zeithamova et al., 2012). While slightly different from the idea of reinstatement of temporal context, described above, these studies establish the critical point that when a stimulus repeats, even in the absence of explicit demands to recall prior encounters, this can trigger spontaneous reinstatement of prior contextual information. Notably, this spontaneous reinstatement of prior contextual information has also been shown to actively prevent subsequent memory interference (Chanales et al., 2019; Kuhl et al., 2010).

These findings also raise other questions. For example, does reinstatement of the original temporal context carry potential costs? In particular, it is possible that reinstatement of the original context interferes with the encoding of the new temporal context (Duncan et al., 2012) or that it increases the probability that the new temporal context will be misattributed to the original temporal context (Gershman et al., 2013; Sederberg et al., 2011). Additionally, another possibility is that reinstatement might subtly bias temporal memory decisions. In other, non-temporal forms of episodic memory, reinstatement of previously-encoded memories that are similar to a current stimulus has been shown to drive differentiation of corresponding MTL representations (Chanales et al., 2017; Favila et al., 2016; Hulbert & Norman, 2015; Kim et al., 2017; Molitor et al., 2021). Given that representational similarity in the MTL has been directly linked to subjective estimates of elapsed time (Ezzyat & Davachi, 2014; Lositsky et al., 2016; Sherman et al., 2023), an intriguing possibility is that reinstatement-induced differentiation of temporal context representations (in the MTL) results in exaggerated memory for the time between stimulus repetitions—a 'repulsive bias.' By this account, reinstatement does not prevent encoding of the new temporal context or induce forgetting of the original context—rather, each temporal context is remembered relatively

accurately, but in a slightly biased manner. As we noted in Introduction, temporal memory is prone to subtle biases (Ezzyat & Davachi, 2014; Lositsky et al., 2016; Poynter, 1983) and repulsive biases in memory have been observed for other, non-temporal features (Chanales et al., 2021; Drascher & Kuhl, 2022; Zhao et al., 2021). Thus, while these ideas are speculative, they represent an interesting avenue for future research that will build a more complete understanding of the cognitive and neural mechanisms supporting temporal memory.

Another important avenue for future work will be to investigate specific ways in which reinstatement may influence subsequent temporal memory. For example, over the past decade, there has been increased interest in the idea that event boundaries can influence cognitive and neural measures of temporal context (DuBrow et al., 2017). In contrast to slowly-drifting temporal context representations, event boundaries can create discontinuities or shifts in temporal context (Clewett et al., 2019; DuBrow, 2024; Radvansky & Zacks, 2017). Event boundaries can take many forms, from salient changes in low-level perceptual information (Heusser et al., 2018) to reward-related prediction errors (Rouhani et al., 2020). Critically, when two stimuli span an event boundary, they are remembered as occurring farther apart in time (Ezzyat & Davachi, 2014; Lositsky et al., 2016; Poynter, 1983), consistent with the idea that event boundaries induce a change in temporal context which is then used to infer the elapsed time. Event boundaries have also been shown to powerfully influence temporal order memory—that is, memory for the relative recency of stimuli (DuBrow & Davachi, 2013; Heusser et al., 2018; Horner et al., 2016; Pu et al., 2022; Rouhani et al., 2020; Wen & Egner, 2022). In particular, order memory tends to be impaired for stimuli that span an event boundary compared to stimuli within the same event. Thus, event boundaries play a key role in shaping temporal context representations and, thereby, influencing temporal memories. As such, a key question to ask is whether event boundaries influence the

degree to which stimulus repetitions induce reinstatement. That is, does an intervening event boundary between repetitions increase, decrease, or not affect the probability that repetition triggers reinstatement of the prior temporal context? Intuitively, it might be expected that a change in context would reduce the probability that repetition would reinstate prior temporal context. However, at least with respect to recognition memory, there are mixed findings with respect to whether a context change between repetitions has any influence on the probability of recognition at S2 (Baddeley, 1982). Interestingly, boundaries themselves have been shown to trigger reinstatement of preceding stimuli (Sols et al., 2017). While speculative, it is possible that if a boundary triggers reinstatement of prior temporal context, this could potentiate additional reinstatement at repetition of stimulus.

References Cited

Baddeley, Alan. D. (1982). Domains of Recollection. *Psychological Review*, *89*(6), 708–729. https://doi.org/10.1037/0033-295X.89.6.708

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

Clewett, D., DuBrow, S., & Davachi, L. (2019). Transcending time in the brain: How event memories are constructed from experience. *Hippocampus*, *29*(3), 162–183. https://doi.org/10.1002/hipo.23074

Davachi, L., & DuBrow, S. (2015). How the hippocampus preserves order: The role of prediction and context. *Trends in Cognitive Sciences*, *19*(2), 92–99. https://doi.org/10.1016/j.tics.2014.12.004

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

DuBrow, S. (2024). Event and Boundaries. In *M. J. Kahana & A. D. Wagner (Eds.), Oxford Handbook of Human Memory* (Vol. 1). Oxford University Press.

DuBrow, S., & Davachi, L. (2013). The influence of context boundaries on memory for the sequential order of events. *Journal of Experimental Psychology: General*, *142*(4), 1277–1286. https://doi.org/10.1037/a0034024

DuBrow, S., Rouhani, N., Niv, Y., & Norman, K. A. (2017). Does mental context drift or shift? *Current Opinion in Behavioral Sciences*, *17*, 141–146. https://doi.org/10.1016/j.cobeha.2017.08.003

Duncan, K., Sadanand, A., & Davachi, L. (2012). Memory's Penumbra: Episodic Memory Decisions Induce Lingering Mnemonic Biases. *Science*, *337*(6093), 485–487. https://doi.org/10.1126/science.1221936

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*(11), Article 11. https://doi.org/10.1038/nrn3827

Ezzyat, Y., & Davachi, L. (2014). Similarity Breeds Proximity: Pattern Similarity within and across Contexts Is Related to Later Mnemonic Judgments of Temporal Proximity. *Neuron*, *81*(5), 1179–1189. https://doi.org/10.1016/j.neuron.2014.01.042

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

Gershman, S. J., Schapiro, A. C., Hupbach, A., & Norman, K. A. (2013). Neural Context Reinstatement Predicts Memory Misattribution. *The Journal of Neuroscience*, *33*(20), 8590–8595. https://doi.org/10.1523/JNEUROSCI.0096-13.2013

Goyal, A., Miller, J., Watrous, A. J., Lee, S. A., Coffey, T., Sperling, M. R., Sharan, A., Worrell, G., Berry, B., Lega, B., Jobst, B. C., Davis, K. A., Inman, C., Sheth, S. A., Wanda, P. A., Ezzyat, Y., Das, S. R., Stein, J., Gorniak, R., & Jacobs, J. (2018). Electrical Stimulation in Hippocampus and Entorhinal Cortex Impairs Spatial and Temporal Memory. *The Journal of Neuroscience*, *38*(19), 4471–4481. https://doi.org/10.1523/JNEUROSCI.3049-17.2018

Heusser, A. C., Ezzyat, Y., Shiff, I., & Davachi, L. (2018). Perceptual boundaries cause mnemonic trade-offs between local boundary processing and across-trial associative binding. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *44*(7), 1075–1090. https://doi.org/10.1037/xlm0000503

Horner, A. J., Bisby, J. A., Wang, A., Bogus, K., & Burgess, N. (2016). The role of spatial boundaries in shaping long-term event representations. *Cognition*, *154*, 151–164. https://doi.org/10.1016/j.cognition.2016.05.013

Huerta, P. T., Sun, L. D., Wilson, M. A., & Tonegawa, S. (2000). Formation of Temporal Memory Requires NMDA Receptors within CA1 Pyramidal Neurons. *Neuron*, *25*(2), 473–480. https://doi.org/10.1016/S0896-6273(00)80909-5

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

Lositsky, O., Chen, J., Toker, D., Honey, C. J., Shvartsman, M., Poppenk, J. L., Hasson, U., & Norman, K. A. (2016). Neural pattern change during encoding of a narrative predicts retrospective duration estimates. *eLife*, *5*, e16070. https://doi.org/10.7554/eLife.16070

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "Time Cells" Bridge the Gap in Memory for Discontiguous Events. *Neuron*, *71*(4), 737–749. https://doi.org/10.1016/j.neuron.2011.07.012

Mau, W., Sullivan, D. W., Kinsky, N. R., Hasselmo, M. E., Howard, M. W., & Eichenbaum, H. (2018). The Same Hippocampal CA1 Population Simultaneously Codes Temporal Information over Multiple Timescales. *Current Biology*, *28*(10), 1499-1508.e4. https://doi.org/10.1016/j.cub.2018.03.051

Molitor, R. J., Sherrill, K. R., Morton, N. W., Miller, A. A., & Preston, A. R. (2021). Memory Reactivation during Learning Simultaneously Promotes Dentate Gyrus/CA $_{2,3}$ Pattern Differentiation and CA $_{1}$ Memory Integration. *The Journal of Neuroscience*, *41*(4), 726–738. https://doi.org/10.1523/JNEUROSCI.0394-20.2020

Montchal, M. E., Reagh, Z. M., & Yassa, M. A. (2019). Precise temporal memories are supported by the lateral entorhinal cortex in humans. *Nature Neuroscience*, *22*(2), 284–288. https://doi.org/10.1038/s41593-018-0303-1

Poynter, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition*, *11*(1), 77–82. https://doi.org/10.3758/BF03197664

Pu, Y., Kong, X.-Z., Ranganath, C., & Melloni, L. (2022). Event boundaries shape temporal organization of memory by resetting temporal context. *Nature Communications*, *13*(1), 622. https://doi.org/10.1038/s41467-022-28216-9

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, *17*, 133–140. https://doi.org/10.1016/j.cobeha.2017.08.006

Rouhani, N., Norman, K. A., Niv, Y., & Bornstein, A. M. (2020). Reward prediction errors create event boundaries in memory. *Cognition*, *203*, 104269. https://doi.org/10.1016/j.cognition.2020.104269

Sederberg, P. B., Gershman, S. J., Polyn, S. M., & Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin & Review*, *18*(3), 455–468. https://doi.org/10.3758/s13423-011-0086-9

Sherman, B. E., DuBrow, S., Winawer, J., & Davachi, L. (2023). Mnemonic Content and Hippocampal Patterns Shape Judgments of Time. *Psychological Science*, *34*(2), 221–237. https://doi.org/10.1177/09567976221129533

Sols, I., DuBrow, S., Davachi, L., & Fuentemilla, L. (2017). Event Boundaries Trigger Rapid Memory Reinstatement of the Prior Events to Promote Their Representation in Long-Term Memory. *Current Biology*, *27*(22), 3499-3504.e4. https://doi.org/10.1016/j.cub.2017.09.057

Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M.-B., & Moser, E. I. (2018). Integrating time from experience in the lateral entorhinal cortex. *Nature*, *561*(7721), 57–62. https://doi.org/10.1038/s41586-018-0459-6

Wen, T., & Egner, T. (2022). Retrieval context determines whether event boundaries impair or enhance temporal order memory. *Cognition*, *225*, 105145. https://doi.org/10.1016/j.cognition.2022.105145

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