

The paradox of prior knowledge: How both predictability and novelty benefit episodic memory

Camille Claire Gasser

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy  
under the Executive Committee  
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

© 2024

Camille Claire Gasser

All Rights Reserved

# **Abstract**

The paradox of prior knowledge: How both predictability and novelty benefit episodic memory

Camille Claire Gasser

We make sense of the present by comparing it to the past. Our capacity for memory makes this comparison possible, characterizing each experience by how much it conforms to or diverges from what we already know. Some events — like commuting to work or cooking your favorite meal — follow a predictable and well-defined structure, whereas others contradict or elude expectations. In this dissertation, I present a series of experiments that explore how the two ends of this continuum, predictability and novelty, affect how we learn from and remember our experiences. Chapter 1 begins by demonstrating that the execution of a predictable and well-learned sequence of actions during learning scaffolds memory for the temporal structure of concurrent events. In Chapter 2, I use functional magnetic resonance imaging (fMRI) to explore the neural basis of this “scaffolding” effect. I find that the brain maintains representations of predictable sequence knowledge during encoding, and that the strength of these representations helps to stabilize activity in the hippocampus and visual cortices and to promote temporal order memory. Finally, Chapter 3 turns to focus on how novelty embedded in our day-to-day lives impacts memory for real-world autobiographical events. Using an intensive longitudinal “daily diary” design, this last study reveals that engaging in new or atypical experiences bolsters memory not only for the novel events in question, but also for non-novel events that occur nearby in time. Taken together, these findings shed light on how episodic memory can benefit from both novelty and familiar structure, illustrating how what we remember is shaped by the expectations we carry with us.

## Table of Contents

List of Figures .....	v
List of Tables .....	vi
Acknowledgements.....	vii
Dedication.....	x
Introduction.....	1
The influence of prior knowledge on episodic memory .....	2
The influence of novelty on episodic memory.....	5
The temporal structure & organization of memory.....	7
Interplay between episodic & motor memory systems .....	10
Investigating memory function in naturalistic contexts .....	12
Overview of the current work .....	13
Chapter 1: Cross-modal facilitation of episodic memory by sequential action execution .....	14
1.1 Introduction .....	15
1.2 Experiment 1 .....	18
1.2.1 Methods .....	18
1.2.2 Results .....	26
1.2.3 Discussion.....	32
1.3 Experiment 2 .....	33
1.3.1 Methods .....	33
1.3.2 Results .....	35
1.3.3 Discussion.....	38
1.4 Experiment 3 .....	39
1.4.1 Methods .....	39
1.4.2 Results .....	41

1.4.3 Discussion.....	45
1.5 General Discussion.....	45
Chapter 2: Predictable action sequences scaffold memory for temporal structure and modulate hippocampal and cortical encoding activity .....	49
B.1 Introduction.....	50
B.2 Methods .....	54
B.2.1 Participants.....	54
B.2.2 Stimuli.....	54
B.2.3 Task design .....	55
B.2.4 MRI acquisition .....	61
B.2.5 fMRI preprocessing .....	62
B.2.6 fMRI analyses .....	65
B.2.7 Statistical analysis.....	70
B.3 Results .....	71
B.3.1 Training performance.....	71
B.3.2 Learning of action sequences persists throughout sequence localizer and encoding tasks .....	72
B.3.3 Execution of predictable action sequences scaffolds temporal order memory.....	73
B.3.4 Object recognition performance .....	74
B.3.5 Neural representation of the predictable action scaffold .....	75
B.3.6 Stability of brain activity within encoding events .....	83
B.3.7 Connecting template activation and within-event stability.....	85
B.4 Discussion.....	89
Chapter 3: Novelty in everyday life promotes long-term memory for autobiographical events.....	97
3.1 Introduction .....	98
3.2 Methods.....	101

3.2.1 Participants .....	101
3.2.2 Study design .....	101
3.2.3 Measuring novelty during the diary period .....	107
3.2.4 Measuring episodic memory for individual events .....	107
3.2.5 Differences between participant groups .....	109
3.2.6 Individual event exclusions .....	110
3.2.7 Statistical analysis.....	110
 3.3 Results .....	111
3.3.1 Behavior during the diary period.....	111
3.3.2 Memory performance .....	113
3.3.3 Effects of novelty on event memory.....	117
3.3.4 Effects of novelty on temporal memory.....	121
 3.4 Discussion .....	123
 Conclusion .....	130
When does predictability help versus hurt memory? .....	131
Effects across the lifespan of a memory.....	134
Remembering the temporal structure of experience .....	136
Implications for memory interventions .....	137
Final remarks.....	138
 References.....	139
Appendix A: Chapter 1 Supplement.....	183
A.1 Supplemental Methods .....	183
A.1.1 Bayesian multilevel regression models .....	183
A.2 Supplemental Results .....	184

A.2.1 Controlling for confounds in the effect of condition on order memory .....	184
A.2.2 Order memory and aisle response accuracy during encoding .....	184
A.2.3 Order memory and aisle repetition during encoding .....	185
A.2.4 Spatial memory performance and aisle location.....	186
A.2.5 Reminder task performance.....	187
A.3 Supplemental Figures .....	188
A.4 Supplemental Tables .....	191
Appendix B: Chapter 2 Supplement .....	193
B.1 Supplemental Methods & Results .....	193
Bayesian multilevel regression models .....	193
Template activation searchlight.....	193
Template activation across sequence positions .....	195
B.2 Supplemental Figures .....	197
B.3 Supplemental Tables.....	199
Appendix C: Chapter 3 Supplement .....	211
C.1 Supplemental Methods & Results .....	211
C.1.1 GPT-4 prompt for detail segmentation .....	211
C.1.2 Supplemental Bayesian regression models.....	212
C.1.3 Characterizing effects across participant groups .....	212
C.1.4 Number of details recalled by event regularity: controlling for diary entry details.....	217
C.1.5 Effects of novelty on memory vividness for non-novel events from adjacent days .....	218
C.2 Supplemental Tables.....	219

## List of Figures

Figure 1.1. Experiment design .....	19
Figure 1.2. Pretraining design and behavior.....	27
Figure 1.3. Order reconstruction memory results for Experiment 1.....	29
Figure 1.4. Spatial memory test performance.....	33
Figure 1.5. Order reconstruction memory results for Experiments 2 and 3 .....	37
Figure 1.6. Item recognition memory test results for Experiment 3.....	44
Figure 2.1. Experiment design.....	57
Figure 2.2. Order memory performance .....	73
Figure 2.3. Encoding pattern similarity across events .....	76
Figure 2.4. Illustration of template activation analysis.....	78
Figure 2.5. Template activation .....	81
Figure 2.6. Within-event stability .....	84
Figure 2.7. Association between template activation and within-event stability .....	87
Figure 3.1. Study design .....	103
Figure 3.2. Overview of event descriptions and variables indexing novelty.....	112
Figure 3.3. Event memory behavior .....	114
Figure 3.4. Temporal memory behavior .....	116
Figure 3.5. Effects of novelty on event memory .....	120
Appendix A, Figure 1. Reaction times (RTs) to aisle cues during pretraining .....	188
Appendix A, Figure 2. Accuracy of aisle responses during encoding.....	189
Appendix A, Figure 3. Item recognition by confidence and sequence position for Experiment 3 .....	190
Appendix B, Figure 1. Template activation whole-brain searchlight.....	197
Appendix B, Figure 2. Template activation across sequence positions .....	198

## List of Tables

Table 3.1. Temporal order accuracy & novelty: model results .....	122
Table 3.2. Subjective temporal distance & novelty: model results .....	123
Appendix A, Table 1. Bayesian model results for Experiment 1 .....	191
Appendix A, Table 2. Bayesian model results for Experiment 2 .....	191
Appendix A, Table 3. Bayesian model results for Experiment 3 .....	192
Appendix A, Table 4. Comparison of memory effects across experiments .....	192
Appendix B, Table 1. Order memory performance: model results (Bayesian) .....	199
Appendix B, Table 2. Item recognition performance: model results (Bayesian) .....	200
Appendix B, Table 3. Template activation & order memory: model results .....	201
Appendix B, Table 4. Within-event stability & order memory: model results .....	202
Appendix B, Table 5. Within-event stability & order memory: model results (Bayesian) .....	204
Appendix B, Table 6. Hippocampal stability & template activation: model results .....	205
Appendix B, Table 7. Hippocampal stability & template activation: model results (Bayesian).....	207
Appendix B, Table 8. LOC stability & template activation: model results.....	207
Appendix B, Table 9. LOC stability & template activation: model results (Bayesian) .....	208
Appendix B, Table 10. Template activation & stability within ROIs: model results.....	209
Appendix B, Table 11. Template activation & stability within ROIs: model results (Bayesian) .....	210
Appendix C, Table 1. Pre-diary and post-diary questionnaire sections .....	219
Appendix C, Table 2. Day-level daily diary variables .....	220
Appendix C, Table 3. Event-level daily diary variables.....	223
Appendix C, Table 4. Memory test questions .....	225
Appendix C, Table 5. Event memory by all novelty variables: model results (Bayesian).....	227
Appendix C, Table 6. Memory vividness by same-day novel event: model results (Bayesian) .....	228

## Acknowledgements

To put into words the love I have for my friends, family, and colleagues feels somehow more difficult than writing this dissertation. Back as an undergrad at NYU, I spent my first few months daunted by the prospect of finding my own community amidst the chaos of this city. Now, over a decade later, I feel safe in the knowledge that I have many — from Greenpoint to Morningside Heights, from New York to California. I appreciate you all so very much.

First: thank you to Lila. Your love for science is a force to behold, and I feel unspeakably grateful to have spent the past 7+ years with you as my mentor. From you, I've learned how to ask good questions, how to believe in the work I produce, and how to find meaning and discovery from an unruly mass of data. I've also learned the value of taking a long walk to sort out my thoughts, and that the greatest part of doing research is sharing ideas with people who challenge and inspire you. I am the scientist I am because of you, and I'll carry that with me no matter what I do next.

Thank you to the members of the Davachi Lab, past and present. To Oded, Emily, Dave, Avi, Alexa, and Lynn: for showing me the ropes, and how much like a family a lab can be. To Nina, Asieh, Amir, Molly, Tarek, Christina, and Ignasi: for your wisdom and guidance, no matter what questions I threw at you. To Natalie and Kathryn: for all you did to keep the lab running and our faces smiling. And of course, to my fellow grad students: Wangjing, John, Victoria, Arlene, and Erin. For all the lunches, the gossip sessions, the ice cream trips, the conferences, and the emotional support. You are all so fiercely intelligent and so relentlessly kind, and I'm so lucky to call you both colleagues and friends.

Thank you also to the rest of my scientific family at Columbia and beyond. To my cohort, Ana, Ben, Anna, Jacob, Margaux, John, and Wangjing. Five years ago, we were gathered in an apartment taking turns playing games on a VR headset. Now, several unprecedent global events later, and we've (almost) made it! What a cool, clever, and inspirational group of people to go on this journey with. To the SIPPS squad, who showed me the joy that comes from advocating for changes we believe in, and who made 'going rogue' so much fun. And to the many others in the psych department who I learned from and

alongside: Hannah, Manasi, Jiawen, Caroline, and many more. Doing a PhD can be a lonely undertaking, full of false starts and pitfalls and moments of doubt. But to work alongside the brilliant people in this department has been a source of comfort even when the light at the end of the tunnel seemed far away.

I'm also so grateful to the faculty who have shared their mentorship and expertise with me throughout the years. To Peter Wais, who gave an unexperienced undergrad the chance to do research for the first time (and had enough trust in me to let me zap participants' brains with a TMS coil). To Shannon Tubridy and Todd Gureckis, for your generous support and patience during my earliest development as a cognitive scientist. Thank you to Chris Baldassano, my secondary advisor: I remember reading your work as a lab manager, stunned by the elegance of the methods you applied and the clarity of the insights they produced. I feel lucky to be counted among those you have mentored. Thank you also to my other committee members — Daphna Shohamy, Deepu Murty, and Theresa Desrochers. I am so excited to share this work with you, and so honored by the opportunity to discuss it together.

To my dearest friends in California: I think the reason I got so homesick when moving to New York was because I had such remarkable people to miss. Rachel: since the moment we became friends after talking about roller coasters in seventh grade, you've been a constant light in my life, an essential source of comfort, and the closest thing to a sister I'll ever have. Kelsey (Matzen): I think you might need to invent another new word to accurately capture how much you mean to me. Thank you for all the nights of singing and dancing, whether at a concert or along the sidewalks of San Francisco. To Gabby, Rachel Jane, Lucy, Corey, Tara, and Max: the older I get, the more acutely I become aware of how lucky I am to have befriended such incredible people so early in my life. Long distance friendships are hard, and one day I'll make it back to the west coast, but in the meantime, thank you for sticking with me regardless.

To my New York family: I am in awe of just how much you all show up for me, how truly cherished I feel by such intelligent, driven, and compassionate people. Emily: the home we've built for ourselves and each other is honestly one of my greatest accomplishments. For all the cozy nights in, the weird and wild nights out; for defending me from bugs and keeping me grounded no matter how hard things got: thank you. Kelsey (Juarez), my bright star, thank you for talking to the girl with red hair on

our first day at NYU. Thank you for the uncountable things you've taught me about what it means to be a friend, and for all the adventures we've had together that I got to be nostalgic for even before they ended. Lauren: what a joy it is to think back on our dinner dates, our questionable decisions and unquestionably fun nights, our early days learning about stats and whimsy under Pascal's guidance. Your empathy, wisdom, and passion for those you love is breathtaking, and I know that I'm better for having known you. To Mariah, thank you for the way you've always made me feel at home, from the West Village to Connecticut to Prague. For all the inside jokes shared over late-night meals, and for the hours spent pouring our hearts out to our favorite bands and to each other. To Olivia: in spite of the frigid air of our Suffolk Street apartment, the warmth I felt that year (and since) when hanging out together — whether playing video games or hitting the town for trivia — has been such a gift. Thank you for driving across New York state to stare at the sun with me, and for all the memories we've yet to make. And to Charlie: on our first date, you made a joke about *p*-hacking, and I fell for you right there. Thank you for helping me breathe in moments of stress, for all the jokes, the games, the reality TV marathons, and for being my best friend. Thank you for pointing out all the good dogs out there, and for introducing me to the wolf himself, Rolo. My love is yours.

Finally, thank you to my family. Preston, I can't even count all the people I've bragged to about you. You are so unbelievably talented, so driven and imaginative, and so, so wonderful to have as a brother. To Grandma Marian: I am endlessly inspired by your curiosity and your devotion to our family. Thank you for all the book recommendations, the phone calls, and the love you've shared with me. To Charlie Brown, the best dog who ever lived, and to Olive and Tootsie, who are giving him a run for his money. And of course: thank you to my mom and dad. This accomplishment is possible because of you — because of your trust in me, your boundless encouragement, your willingness to let me fall apart and your ability to build me back up afterward. To list all the ways in which you've both inspired me would add another hundred pages to this dissertation, and it would still fail to convey how grateful I am for all that you are and all that you do. So instead I'll just say: thank you for the way you see me, for the love you've given me, and for being the best parents anyone could ask for. I love you.

## **Dedication**

To Grandma Verna, Grandpa Lee, and Grandpa Jack.

## Introduction

Early this year, I found myself standing amidst the vast expanse of New York City's Javits Convention Center. My purpose: to attend the American Kennel Club's annual "Meet the Breeds" event, where enthusiastic dog-lovers congregate to learn about — and pet! — dogs of over a hundred different breeds. While there, I met some dogs I'd never seen before: a lumbering giant with fur the texture of an unruly mop; an all-white spaniel with a short nose and shorter legs. And I met dogs easily recognizable from past experience: a playful golden retriever; an impossibly fluffy Saint Bernard.

Despite the silliness of this anecdote, it highlights an important quality of our experiences: that they range along a continuum from the very familiar to the completely new. This dimension of experience, moreover, shapes how we learn and what we remember. An event that deviates from our existing memories can signal an important learning opportunity, giving us the chance to update our understanding of the world with new and useful information. Events that conform to past experiences, conversely, reinforce our expectations and allow us to draw connections between related situations, a process critical to the construction of abstract and generalizable knowledge. Both kinds of events are important to remember, and both tap into unique neurocognitive learning mechanisms.

In this dissertation, I draw inspiration from multiple lines of previous research to examine how episodic memory (i.e., memory for specific past events) is shaped by both the prior knowledge we possess and the novelty of our experiences. In doing so, I focus heavily on how prior knowledge and novelty influence memory for temporal structure, or the way that an experience unfolds across time. I begin by conducting an in-depth investigation into how prior sequential knowledge — more specifically, a well-learned sequence of motor actions — can scaffold temporal memory for novel episodic events (Chapters 1 & 2). Across multiple experiments, I examine both the behavioral and neural mechanisms that drive this memory facilitation. Then, I explore how novelty and typicality in everyday life impact memory for real-world autobiographical experiences (Chapter 3). To provide the necessary context for these research aims, this introduction will briefly review several relevant areas of prior literature.

## The influence of prior knowledge on episodic memory

Many situations we face bare resemblance to things we've encountered in the past. These commonalities across experiences allow us to extract complex and generalized knowledge about how the world works: how to navigate an airport, how to resolve conflict with a loved one, or what to expect on the first day of a new class. The creation and function of prior knowledge structures has been of significant interest to researchers over the last several decades (Alba & Hasher, 1983; J. R. Anderson, 1981; Bartlett, 1932; Brod et al., 2013; Farzanfar et al., 2023; Ghosh & Gilboa, 2014; van Kesteren et al., 2012). In particular, many have explored the question of how the prior knowledge we possess shapes our ability to learn novel information. In one early study addressing this question, Anderson et al. (1978) presented participants with a set of food words that were embedded within one of two stories: one describing a meal in a restaurant, and one describing a supermarket trip. Memory for food items was enhanced in those who read the restaurant story, ostensibly because participants could leverage their schematic knowledge about the content and structure of a typical meal to "scaffold" novel learning. Similar results have since been observed across a wide variety of experimental paradigms. To name several examples: those with more expertise in a particular subject are better able to acquire novel facts related to that topic (Overschelde & Healy, 2001; Rawson & Van Overschelde, 2008; van Kesteren et al., 2014; Woollett & Maguire, 2010); novel associations that include famous people or places are easier to remember than those that do not (Bein et al., 2019; Bellana et al., 2021; Z. X. Liu et al., 2017; Reder et al., 2013); and embedding items within a familiar spatial environments enhances learning in both animals (Cox et al., 2017; Tse et al., 2007, 2011) and humans (Guo & Yang, 2020, 2023; Sommer, 2017; van Buuren et al., 2014). However, alongside these benefits, prior knowledge can also distort or bias memory behavior, such as by increasing false memories for schema-congruent information (Bellana et al., 2021; Kleider et al., 2008; Lew & Howe, 2017; Umanath & Marsh, 2014).

As suggested above, prior knowledge is thought to aid memory by providing an organizational scaffold for incoming information (J. R. Anderson, 1981; Ghosh & Gilboa, 2014). In this way, new

memory traces can become rapidly integrated with existing memory representations (McClelland, 2013), potentially strengthening their durability and/or increasing their accessibility during later retrieval.

Although this framework is generally endorsed throughout the field, and is supported by neuroscience findings (see below), other plausible mechanisms have been proposed for this memory benefit. For example, it has been argued that prior knowledge during learning frees up attentional resources for the elements of one's experience that are new, bolstering their strength in episodic memory (DeWitt et al., 2012; Diana & Reder, 2006; Reder et al., 2016). In Chapter 1 of this dissertation, I consider both of these mechanisms — scaffolding versus attentional resources — while investigating how prior sequential knowledge impacts memory for concurrent episodic events.

***Neural mechanisms of prior knowledge-mediated memory.*** Research points to a special role for the medial prefrontal cortex (mPFC) in representing prior knowledge structures and mediating their influence on novel learning (Gilboa & Marlatte, 2017; Preston & Eichenbaum, 2013; van Kesteren et al., 2012). Lesion studies, for example, have revealed that patients with ventral mPFC damage tend to be *less* susceptible to false memories emerging from semantic knowledge (Melo et al., 1999; Warren et al., 2014), and do not show typical memory enhancements for schema-congruent associations (Spalding et al., 2015). In healthy humans, mPFC activation has also consistently been observed when participants encode information that is congruent with prior knowledge (Bonasia et al., 2018; Brod & Shing, 2018; Guo & Yang, 2023; Z. X. Liu et al., 2017; Raykov et al., 2020; van Kesteren, Beul, et al., 2013; van Kesteren et al., 2014, 2020), as well as when they retrieve this information later on (Brod et al., 2015; Guo et al., 2023; van Kesteren, Rijpkema, et al., 2010; Wagner et al., 2015). Modern neuroimaging analysis methods (e.g., representational similarity analysis, Kriegeskorte et al., 2008) have also made it possible to probe the informational content reflected by brain activity. Such approaches have provided further support for the claim that mPFC contains representations of complex knowledge structures, including real-world concepts or event schemas (Audrain & McAndrews, 2022; Baldassano et al., 2017; Masís-Obando et al., 2022), knowledge about people or characters (Raykov et al., 2021), rule or concept knowledge (Mack et al., 2020; Schuck et al., 2016; Sommer et al., 2022), and familiar spatial layouts (Zheng et al., 2021).

After initial learning, episodic memories are thought to undergo the process of systems consolidation, whereby novel engrams initially dependent on the hippocampus become gradually distributed throughout the neocortex (Dudai et al., 2015; Nadel & Moscovitch, 2000; Winocur & Moscovitch, 2011). The presence of prior knowledge during learning can accelerate this consolidation process (Gilboa & Marlatte, 2017; McClelland, 2013; Morris, 2006; Sommer et al., 2022; van Kesteren et al., 2012) — consistent with the idea discussed above that knowledge-congruent information benefits from rapid integration with existing memories. In a classic study supporting this hypothesis, rodents received hippocampal lesions shortly after learning novel associations that were either consistent or inconsistent with a learned spatial schema; whereas memory for schema-inconsistent associations was lost after lesioning, schema-consistent associations were retained, suggesting that these memory traces had quickly become independent of the hippocampus (Tse et al., 2007). Follow-up experiments then indicated that schema-consistent learning was associated with upregulation of plasticity-related gene expression in the rodent analogues to mPFC and retrosplenial cortex (Tse et al., 2011).

Human imaging studies provide complementary evidence for altered interactions between the hippocampus and mPFC in prior knowledge-mediated learning. Interestingly, evidence for the directionality of this relationship is mixed, with some studies finding support for enhanced mPFC-hippocampal connectivity when novel episodic information is compatible with prior knowledge (Audrain & McAndrews, 2022; Z. X. Liu et al., 2017; Sommer, 2017), and others finding the opposite pattern (Bein et al., 2014; Bonasia et al., 2018; van Kesteren et al., 2014; van Kesteren, Fernandez, et al., 2010). These contradictory findings may be driven by many factors, including the relationship between prior knowledge and new content (e.g., Brod & Shing, 2018), the hippocampal subregion under consideration (e.g., Guo et al., 2023; Guo & Yang, 2020), and the timepoint in the memory cycle at which brain activity is examined (i.e., encoding, retrieval, or consolidation). Nevertheless, such work collectively points to the involvement of mPFC, the hippocampus, and their interactions when new learning occurs in the context of prior knowledge. It is also important to note that brain regions beyond mPFC have been implicated in both the representation of prior knowledge and this hippocampal-neocortical dialogue. For example,

schema representations have been observed throughout the default mode network, including in posterior medial cortex (PMC) and lateral parietal cortex (Baldassano et al., 2017; Masís-Obando et al., 2022). PMC is also frequently engaged (alongside mPFC) during the encoding (Bonasia et al., 2018; Maguire et al., 1999; Sommer, 2017; Sommer et al., 2022) and retrieval (Sommer, 2017; Sommer et al., 2022; van Buuren et al., 2014; Wagner et al., 2015) of knowledge-consistent information. As part of Chapter 2, I explore how both mPFC and PMC represent learned sequential knowledge, as well as how these regions interact with the hippocampus to support novel encoding.

## The influence of novelty on episodic memory

Alongside the rich history of work on how existing knowledge can scaffold new learning, a similar breadth of research speaks to the mnemonic benefits of novelty (Duszkiewicz et al., 2019; Kafkas & Montaldi, 2018; Lorents et al., 2023; Ranganath & Rainer, 2003; Schomaker & Meeter, 2015). Many of our most vivid memories correspond to moments that are new or unexpected: getting married, seeing a solar eclipse, or taking a trip to a far-away country. From an evolutionary perspective, this novelty benefit makes sense: in order to adapt to a complex and dynamic environment, we need to be capable of detecting information we lack experience with and storing it for future use.

The effects of novelty on episodic memory have been examined via multiple kinds of experimental paradigms. Several studies, for example, have assessed the learning of stimuli that are novel (versus familiar) within the confines of a single experimental session. Much of this work finds that item memory for these novel stimuli is better than for more familiar stimuli (Kormi-Nouri et al., 2005; Tulving & Kroll, 1995; but see also Kafkas & Montaldi, 2015; Poppenk et al., 2010). Novelty can also be induced by presenting participants with information that violates their learned expectations. These unexpected events elicit increased attention (Berlyne, 1951; Friedman et al., 2001; Horstmann, 2002; Reisenzein et al., 2019), and are often remembered better than events that conform to predictions (Hunt, 1995; Kafkas & Montaldi, 2015; Ranganath & Rainer, 2003; Schomaker & Meeter, 2015). Such findings can be interpreted through the lens of a predictive framework, in that an unexpected stimulus elicits a memory

“prediction error” that signals a need for the brain to update its internal model of the world (Frank & Kafkas, 2021; Henson & Gagnepain, 2010; Reichardt et al., 2020). These forms of mnemonic prediction errors, indeed, have been shown to enhance multiple forms of learning (Den Ouden et al., 2012; Garrison et al., 2013; Rouhani & Niv, 2021), including episodic memory (Bein et al., 2021; Ergo et al., 2020; A. Greve et al., 2017; Rouhani et al., 2018).

***Novelty processing in the brain.*** The detection of novelty evokes a cascade of neurobiological signals that shape memory encoding by modulating hippocampal function (Kafkas & Montaldi, 2018). The route by which this modulation takes place may depend on the type of novelty in question. According to one popular theoretical model (Kafkas & Montaldi, 2018), “contextual novelty” — which arises when expectations are violated — elicits increased coupling between the hippocampus and the dopaminergic midbrain, leading to a release of dopamine within the hippocampus that enhances learning by boosting long-term potentiation, or LTP (Lisman & Grace, 2005; Shohamy & Adcock, 2010). “Absolute novelty,” triggered by a stimulus or event that is wholly unfamiliar, instead involves the release of acetylcholine, a neurotransmitter that has also been linked to long-term memory (Blokland, 1995; Gold, 2003; Grön et al., 2005; Hasselmo, 2006). This framework (along with other related ideas, e.g., Duszkiewicz et al., 2019) thus identifies multiple neuromodulatory pathways by which novelty can bolster learning.

***Novelty’s penumbra.*** So far, I have reviewed research indicating that episodic memory is enhanced for novel information. Interestingly, an additional body of work establishes that engaging in a new experience can enhance learning during a *separate* task or trial that occurs close by in time (Duncan et al., 2012; Lorents et al., 2023; Schomaker & Meeter, 2015). In rodents, brief exploration of an unfamiliar spatial environment shortly before *or* after an unrelated learning task has been shown to boost LTP and facilitate memory for that task (Ballarini et al., 2009; Davis et al., 2004; Li et al., 2003; Moncada & Viola, 2007). In humans, several experiments conducted in classroom settings have reported similar effects: students who engaged in a novel experience (e.g., a new science lesson) within close temporal proximity to a separate learning task performed better on that task (Ballarini et al., 2013) — even when memory was tested weeks later (Ramirez Butavand et al., 2020). Similar proactive and retroactive

memory enhancements can also be triggered by moments of high arousal or emotion (Dunsmoor et al., 2015; McGaugh, 2018), responses that often co-occur with novelty. These findings have been taken collectively as support for a novelty-triggered “behavioral tagging” or “tag-and-capture” mechanism, whereby a salient event triggers the release of plasticity-related proteins that potentiate weak memory traces (Dunsmoor et al., 2022; Moncada et al., 2015; Redondo & Morris, 2011). Interestingly, research has also shown that such potentiation can operate selectively, such that a highly novel or arousing experience specifically rescues memories that have a meaningful connection to that experience (e.g., shared semantic information; Dunsmoor et al., 2015; Patil et al., 2017). Ultimately, this body of work speaks to the importance of understanding how the novelty embedded in our experiences shapes what we remember — a question I explore in a naturalistic context throughout Chapter 3.

## The temporal structure & organization of memory

Several decades ago, Endel Tulving described episodic memory as a form of “mental time travel” (Tulving, 1985), speaking to the profound ability of our memories to transport our awareness to distant times and places. This description highlights the fundamental importance of temporal information to what and how we remember. Temporal structure is ubiquitous in everyday acts of memory retrieval: ask your friend what they did today, for example, and they are likely to guide you through a timeline of their recent activities. The “temporal contiguity effect” — whereby items encoded close together in time are recalled successively during memory retrieval — is one of the most robust and well-documented effects in memory research (Healey et al., 2019; Kahana et al., 2008). By allowing us to learn relationships between cause and effect, temporal memory also underpins our ability to generate predictions about the future — arguably one of the most central functions of the brain (Bar, 2007; Bubic et al., 2010; Buckner & Carroll, 2007; Spratling, 2016). Investigating how we remember the temporal structure of experience, therefore, is critical to the goal of understanding episodic memory.

**Temporal coding in the brain.** The hippocampus plays a central role in representing temporal information (Eichenbaum, 2014, 2017; Howard & Eichenbaum, 2013; Ranganath & Hsieh, 2016).

Damage to the hippocampus impairs memory for sequential order, both in animals (DeVito & Eichenbaum, 2011; Farovik et al., 2010; Fortin et al., 2002; Kesner et al., 2002) and humans (DuBrow et al., 2024; Mayes et al., 2001). Neuroimaging studies consistently report hippocampal activation during the encoding and retrieval of order information (Lehn et al., 2009; Schendan et al., 2003; Tubridy & Davachi, 2011), as well as evidence that hippocampal activity patterns represent temporal structure in the environment (Bellmund et al., 2022; Hsieh et al., 2014; Kyle et al., 2015; C. Liu et al., 2021; Nielson et al., 2015). The hippocampus also contains a special population of neurons categorized as “time cells,” which fire selectively at specific moments within an interval of time (MacDonald et al., 2013; Manns et al., 2007; Shahbaba et al., 2022). Of note, sensitivity to temporal or sequential information is not exclusive to the hippocampus, with other work pointing to important roles for the nearby lateral entorhinal cortex (Bellmund et al., 2019, 2020; Hasselmo & Stern, 2014; Tsao et al., 2018) and the prefrontal cortices (Desrochers et al., 2015, 2019; Fuster, 2001).

***Representation of temporal context.*** The existence of hippocampal time cells and other temporally sensitive neural representations fits broadly with the Temporal Context Model (TCM) of episodic memory. This framework argues for the existence of an internal temporal context signature that changes gradually across time (Howard, 2017; Howard & Kahana, 2002; Polyn et al., 2009; Polyn & Kahana, 2008). This context representation integrates one’s current perceptual input with a temporally graded summation of previous inputs, such that items encoded close by in time will be associated with more similar temporal context representations than those encoded at more distant moments. Mathematical formulations of TCM and its successors can successfully account for many aspects of memory behavior, including temporal contiguity during recall and recency effects (Sederberg et al., 2008).

Importantly, representations of temporal context are also shaped by the overarching structure of our experience. Although time unfolds continuously, we tend to remember our lives as a sequence of discrete episodes or events (e.g., having breakfast, commuting to work). Theories of “event segmentation” posit that these episodes arise from stability and change in one’s surrounding context (Clewett et al., 2019; Clewett & Davachi, 2017; DuBrow et al., 2017; Kurby & Zacks, 2008): experiences that take place

in the same context (e.g., because they occur in the same spatial environment, while one is engaged in the same task, etc.) tend to be integrated together in memory, whereas experiences from different contexts are more separated — even when they occur close in time. This idea is supported by several robust behavioral findings. Using the well-validated Ezzyat-DuBrow-Davachi (EDD) paradigm (Buonomano et al., 2023), it has been shown that participants have better memory for the order of items that belong to the same event than those that bridge across an event boundary (i.e., the transition between one event/context and the next), suggesting that items associated with the same context are more strongly integrated (Clewett, Gasser, et al., 2020; DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011; Heusser et al., 2018; Horner et al., 2016). In addition, pairs of stimuli from the same context are often remembered as occurring closer together in time than those from distinct contexts — even when controlling for the objective amount of time elapsed (Brunec et al., 2020; Clewett, Gasser, et al., 2020; Ezzyat & Davachi, 2014).

In the brain, the temporal structure of experience may be reflected in part by the stability of neural activity across time. In studies of narrative perception, for example, neural activity patterns throughout the default mode network remain fairly stable within a given context (e.g., a single scene in a movie), but have been reported to change rapidly when an event boundary is perceived (Baldassano et al., 2017; C. S. Lee et al., 2021; Silva et al., 2019). However, open questions remain with respect to how exactly stability supports temporal or sequential memory behavior. In studies focused on the hippocampus, some findings indicate that temporal integration is associated with greater similarity (i.e., stability) across hippocampal representations during encoding (DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014; Kyle et al., 2015). Others, however, have found that greater *dissimilarity* in hippocampal patterns across time supports order memory (Jenkins & Ranganath, 2016). Whether temporal memory benefits more strongly from neural integration or differentiation may depend on multiple factors, including how memory is tested, what strategies an individual uses during encoding, or what brain regions are under examination (Bein & Davachi, 2024; Bellmund et al., 2022; DuBrow & Davachi, 2017).

Ultimately, temporal structure is a fundamental property of episodic memory, influencing both the organization of what we remember and the way that our experience is represented in the brain. In my

own research, I seek to build upon the effort made thus far to characterize the mechanisms of temporal memory by asking how these mechanisms are shaped by prior knowledge and novelty. In particular, throughout Chapters 1 and 2, I investigate how existing knowledge about sequential motor behavior facilitates temporal integration during novel episodic events — potentially by stabilizing brain activity across time. In Chapter 3, I then ask how the novelty embedded within our experience impacts how we remember the order of and distance between autobiographical events.

## **Interplay between episodic & motor memory systems**

As just mentioned, in Chapters 1 and 2 of this dissertation, I use learned motor behavior as a vehicle to explore how prior sequential knowledge shapes temporal memory for novel experiences. More specifically, I ask participants to encode novel events while simultaneously executing sequences of motor actions that are either well-learned or unpredictable. In addition to investigating the effects of sequential knowledge structures, such work bridges across the domains of episodic memory and motor learning, areas of neuroscience research that have largely progressed in isolation. This separation exists in part due to the long-standing characterization of episodic and motor memory as emerging from distinct memory systems with different neural substrates.

***Multiple memory systems.*** Memory has been historically divided into multiple systems that fall under two broad categories: declarative and non-declarative. Declarative memory involves the conscious recollection of facts or events and encompasses both episodic and semantic memory systems; non-declarative memory refers to forms of learning that are more implicit in nature, such as procedural memory (i.e., for skills or habits), priming, and conditioning (Squire & Zola, 1996). Compelling evidence for a division between memory systems comes from studies conducted in patients with brain damage, which have revealed that selective impairment of one type of memory can occur in the absence of deficits on other learning tasks (Hopkins et al., 2004; Reber & Squire, 1994). For example, patient H.M. — who famously lost the capacity to form new episodic memories after undergoing a bilateral hippocampal lobectomy — was nevertheless able to acquire new motor skills (Corkin, 1968; Milner et al., 1968). In the

years following this seminal work, findings amassed to support the idea that episodic memory capacity is rooted in the medial temporal lobe, which includes the hippocampus as well as nearby parahippocampal, perirhinal, and entorhinal cortices (Squire & Zola, 1996; Tulving & Markowitsch, 1998). In contrast, motor learning — typically classified as a form of procedural memory — recruits the striatum and cerebellum, along with a network of sensorimotor and prefrontal regions (Exner et al., 2002; Graybiel & Grafton, 2015; Yokoi & Diedrichsen, 2019).

However, recent work suggests that the division between memory systems is less stark than previously characterized. Despite its integral role in episodic memory, the hippocampus has also been shown to contribute to motor sequence memory (Albouy et al., 2008, 2015; Curran, 1997; Jacobacci et al., 2020; King et al., 2022; Schapiro et al., 2019; Schendan et al., 2003) and to statistical learning more broadly (Covington et al., 2018; Henin et al., 2021; Schapiro et al., 2014, 2017). A recent review from Sherman et al. (2024) specifically advocated for an adjustment of the multiple memory systems taxonomy, highlighting evidence that brain regions thought to be selectively involved in one type of memory (e.g., the hippocampus) are capable of supporting more varied computations (also see Shohamy & Turk-Browne, 2013), and that interactions between subsystems are common (e.g., Shohamy & Wagner, 2008). Taken together, these findings allow for the possibility that different forms of memory can work together in support of learning.

***Evidence for cross-modal memory interactions.*** Although behavioral evidence for interactions between motor and episodic memory is limited, some promising findings exist. Paradigms examining “enacted encoding,” for example, show that engaging in congruent motor actions during novel episodic encoding (e.g., making a scrubbing motion when presented with the word “clean”) improves subsequent item memory, relative to perception of the stimuli alone (Engelkamp, 1998; Engelkamp & Cohen, 1991; Madan & Singhal, 2012; Noroozian et al., 2022). Although most work on enacted encoding considers cases where there is a meaningful connection between the stimulus and action, some studies show similar memory enhancements even when actions are unrelated to episodic content (Kinder & Buss, 2021; Yebra et al., 2019). This suggests that motor actions can and do influence concurrent memory processes, even in

the absence of meaningful associations between actions and novel stimuli.

Critically, however, this existing work focuses predominantly on how actions impact item memory (e.g., memory for individual words or images). In contrast, as discussed in the previous section, we often remember our experiences as cohesive events, each composed of an ordered sequence of information (Clewett et al., 2019; Clewett & Davachi, 2017). Some of the only attempts to examine motor-memory interactions during learning of sequential content have used paradigms where motor sequences and episodic content are encountered *asynchronously*. These studies show that motor sequence learning during one task can “transfer” to memory performance on a separate, word-list learning task (and vice versa) (Moshé & Robertson, 2016; Robertson, 2022). While such findings reinforce the feasibility of identifying cooperative effects between motor behavior and memory, how motor sequences impact concurrent encoding of episodic content remains an open question.

## **Investigating memory function in naturalistic contexts**

In recent years, there has been a dramatic increase in studies employing “naturalistic” paradigms to investigate memory function (Pooja et al., 2024; Shamay-Tsoory & Mendelsohn, 2019; Virk et al., 2024). These efforts have adopted many different approaches, such as using complex narrative stimuli (Baldassano et al., 2017; Dev et al., 2022; Jääskeläinen et al., 2021), leveraging virtual reality tools or video games (La Corte et al., 2019; Plancher et al., 2013; Reggente et al., 2020), having participants encode information in semi-controlled immersive settings (Cliver et al., 2024; Diamond & Levine, 2020; Jeunehomme et al., 2022; Stasiak et al., 2023), and quantifying autobiographical memories via retrospective reports (Addis et al., 2004; Bonnici et al., 2012; Levine et al., 2002; Robinson, 1976) or experience sampling methods (Chow & Rissman, 2017; Doherty et al., 2012; Martin et al., 2022; Rasmussen et al., 2014). Although these paradigms carry their own challenges — and certainly do not replace the advantages of well-controlled laboratory experiments — they underscore the importance of studying how we remember experiences that are formed in more complex and dynamic environments.

In my own work, I meet this goal in two ways. Throughout Chapters 1 and 2, I investigate how

active behavior — something that is ubiquitous in daily life but often treated as a confound in memory experiments — influences the concurrent formation of episodic memories. Then in Chapter 3, I collect written records of participants’ real-life experiences to probe their personal memories, making it possible to examine the relationship between novelty and memory in a truly naturalistic setting.

## **Overview of the current work**

To summarize, in this dissertation, I use a combination of behavioral and neuroimaging methods to investigate how both prior knowledge and novelty affect the structure and content of episodic memory. Chapter 1 introduces an original experimental paradigm that explores how prior sequential knowledge can scaffold memory for new episodic events. Across three experiments, I find that engaging in a well-learned sequence of motor actions during the encoding of novel item sequences selectively enhances temporal order memory for those items. Chapter 2 then expands upon these findings by using fMRI to measure brain activity as participants engage in this sequence encoding task. This study reveals that learned motor sequence knowledge is represented in a distributed network throughout the brain, encompassing mPFC, PMC, as well as motor-selective regions. Moreover, the fidelity of these representations during encoding events is associated with the stability of neural activity in the hippocampus across time, which in turn is linked to successful temporal order memory. Finally, Chapter 3 presents an intensive longitudinal “daily diary” study that explores the link between autobiographical memory and the extent to which participants engage in novel versus typical experiences from day to day. Here, we show that greater experiential novelty has beneficial and distributed effects on memory for autobiographical events, but does not strongly impact memory for the temporal relationships between them.

# **Chapter 1: Cross-modal facilitation of episodic memory by sequential action execution**

Camille Gasser & Lila Davachi

Please note, chapter published as:

Gasser, C., & Davachi, L. (2023). Cross-modal facilitation of episodic memory by sequential action execution. *Psychological Science*, 34(5), 581-602.

## 1.1 Introduction

Our lives are imbued with structure and regularity. In particular, many of our behaviors can be organized into repetitive, well-defined sequences of actions. Consider your commute: you might walk to the subway station, then swipe your ticket at the turnstile, then step aboard the train. While engaging in this familiar sequence of actions, you are also likely to encounter an influx of novel sensory information, which may have little to do with your actual behavior (e.g., receiving a text while entering the subway). Despite the prevalence of familiar action sequences in everyday life, we know little about their impact on memory for novel experiences that unfold simultaneously. The current study was therefore designed to explore this cross-modal interplay between familiar action execution and novel episodic memories.

This investigation bridges two fields — motor learning and episodic memory — that have evolved in parallel with relatively little cross-talk. Existing work at this intersection reveals that engaging in actions during learning boosts memory relative to passive encoding (Engelkamp & Cohen, 1991; Yebra et al., 2019), and that learning can transfer between motor sequences and other types of information (e.g., repeatedly-encoded word lists) (Mosha & Robertson, 2016; Mutanen et al., 2020; Thibault et al., 2021). Separately, a line of research on stimulus-response associations has characterized how motor actions can become rapidly integrated with the stimuli they are paired with, which serves to facilitate processing of those stimuli or actions during future encounters (Dobbins et al., 2004; Hommel, 2004; Schnyer et al., 2006). Through this binding mechanism, stimuli can be linked to individual actions even after a single stimulus-response encounter. However, these previous lines of work do not directly address how executing sequences of familiar actions impacts the simultaneous encoding of novel information. Furthermore, an understanding of how sequential actions affect multiple forms of one-shot episodic memory is lacking.

In previous literature, it has been well-established that prior knowledge about the world can facilitate novel learning (e.g., Alba & Hasher, 1983; Anderson et al., 1978; Bransford et al., 1972; Kole & Healy, 2007; Tse et al., 2007; van Kesteren et al., 2010, 2014). These findings even form the basis of

strategic techniques like the Method of Loci, in which people boost memory for new information by deliberately associating it with familiar spatial contexts (Bass & Oswald, 2014; McCabe, 2015; Reggente et al., 2020). However, the lion’s share of this existing work focuses on how prior knowledge supports memory for information that belongs to the same domain or is meaningfully related to that knowledge. Here, we ask instead how familiar motor actions impact memory for novel, unrelated episodic experiences that belong to a different modality than those actions.

We propose that familiar action sequences should enhance memory for concurrent experiences by providing a “temporal scaffold” for incoming information. When one executes a well-known motor sequence, the representation of that sequence may become activated in memory and remain online throughout the event. As new information is encountered, it can then become integrated within (or “slotted into”) the existing event representation through their co-activation. This “scaffolding hypothesis” makes specific predictions for how action sequences should promote novel learning, based on multiple avenues of previous work. In particular, the hippocampus has been shown to support memory for both motor and episodic item sequences (Albouy et al., 2008, 2015; Curran, 1997; Schendan et al., 2003), and to form stable neural representations of sequential information that is studied repeatedly (Hsieh et al., 2014; Kalm et al., 2013; Paz et al., 2010). Greater stability in hippocampal activity patterns, in turn, can enhance temporal binding between items (DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014). Integrating these results, we hypothesize that when participants encode new information during execution of a known action sequence, these items become linked to a highly stable event representation, which ultimately improves temporal memory.

This putative mechanism draws inspiration from research on event segmentation, which posits that stability and change in one’s surrounding context help to organize continuous experience into discrete, unified memory episodes (Clewett & Davachi, 2017; Davachi & DuBrow, 2015). While some work in this space has focused on how temporal memory is negatively impacted by boundaries *between* events (e.g., abrupt shifts in spatial context or task goals; Clewett et al., 2020; Horner et al., 2016; Kurby & Zacks, 2008), here we explore a mechanism by which memory for information from *within the same*

*event* can be better integrated. Engaging in a familiar behavioral sequence may afford individuals with a level of mental context stability that benefits within-event temporal binding beyond what is granted by consistency in perceptual input or task demands alone. Our scaffolding hypothesis is also broadly consistent with extant models of prior knowledge-mediated learning, which argue that enhanced memory for new information occurs because these items are assimilated into existing memory representations through hippocampal-neocortical connectivity (McClelland, 2013; Gilboa & Marlatte, 2017).

We also test the alternative hypothesis that familiar behaviors boost memory by freeing attentional resources. That is, when some part of our environment is familiar, we can direct more focus to information that is new (DeWitt et al., 2012; Reder et al., 2016). Under this “attentional resource hypothesis”, memory for all novel elements of an experience — not just temporal order — should be enhanced by familiar action execution.

Critically, both our scaffolding and attentional resource hypotheses raise the question of how motor actions *per se* drive any resulting memory improvements. It could be, for example, that engaging in familiar actions improves memory simply because individuals have predictions or expectations about the future. In this case, the capacity to make memory-based predictions about *any* element of upcoming events may be sufficient to scaffold concurrent learning. Existing work shows that the relationship between prediction and novel encoding is not clear-cut (Ritvo et al., 2019; Sherman & Turk-Browne, 2020). Nevertheless, we acknowledge this possibility, and here have chosen to focus on motor actions as the modality through which such predictions are formed and expressed. This decision is motivated both by the prevalence of familiar behaviors in our day-to-day lives, and by the previously-discussed research on demonstrated interactions between motor and episodic memories (e.g., Engelkamp & Cohen, 1991; Mosha & Robertson, 2016). Further, this study will help to clarify and explore the range of circumstances in which prior knowledge and prediction aid novel learning.

Therefore, to investigate whether and how familiar action sequences impact new learning, we designed a novel paradigm (Figure 1.1) in which participants encoded sequences of novel items while embarking on “errands” to two different stores. During each errand, they were also required to execute a

sequence of simple motor actions. Critically, in one of the two stores — the *predictable store* — the sequence of actions executed during encoding always followed a familiar, well-learned pattern. We then examined whether temporal memory for the order of novel items encoded within the predictable store was enhanced relative to memory for novel items encountered in the *random store*, where actions were unpredictable. Further, we adjudicate between the scaffolding and attentional resources hypotheses by asking whether familiar action sequences benefit memory only for the order of novel items, or also for their perceptual details and spatial contexts. Across three experiments, we find that executing learned action sequences robustly and selectively bolsters memory for the temporal sequencing of novel information, consistent with our scaffolding hypothesis.

### Open Practices Statement

All data, stimuli, and analysis scripts for this study are publicly available via OSF and can be accessed at <https://osf.io/xgwzf/>. The experiments reported in this article were not preregistered.

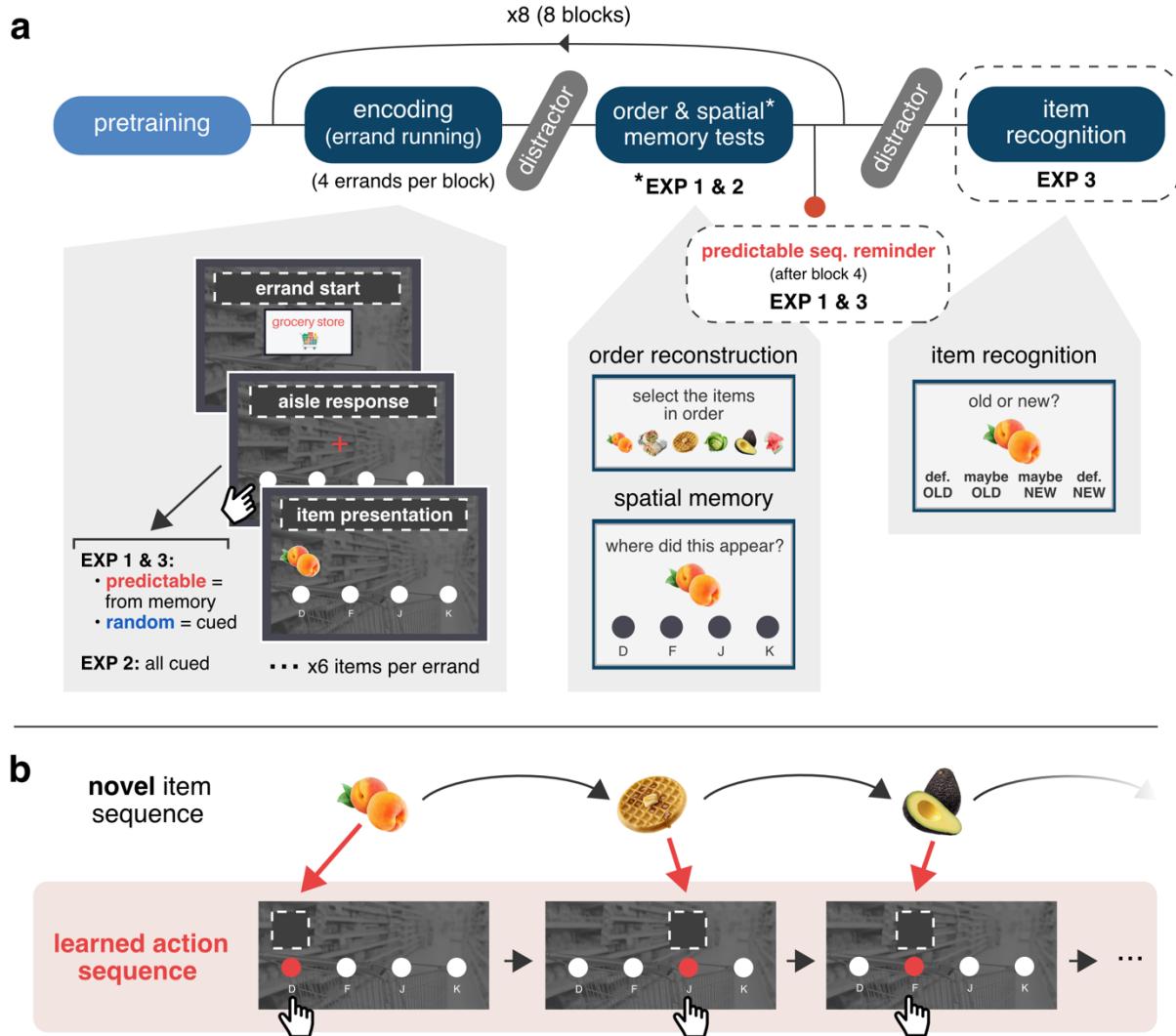
## 1.2 Experiment 1

### 1.2.1 Methods

#### *Participants*

A total of  $N = 80$  participants (29 female) were recruited online via Prolific ([www.prolific.co](http://www.prolific.co)). Previous studies in our lab have used samples in the range of 25-35 participants when exploring other effects of event structure on episodic memory (e.g., Clewett et al., 2020; DuBrow & Davachi, 2013). Given that here we explored a novel behavioral effect with an unknown size, and that our data collection involved online participants (who likely exhibit greater heterogeneity than those recruited on university campuses), we chose to enroll a considerably larger sample size to ensure sufficient statistical power. Specifically, a sample of  $N = 80$  gives us 80% power to detect effect sizes with at least size  $d = 0.31$  (i.e., small-medium effects). All participants were healthy adults between the ages of 18-35 ( $M = 24.9$ ,  $SD = 4.9$ ), reported having normal or corrected-to-normal vision, and had a Prolific study approval rate of at least 60%. Participants received \$14 upon study completion (duration  $M = 61.0$  min,  $SD = 19.9$ ). All

participants provided informed consent, and all procedures were approved by the Institutional Review Board at Columbia University.



**Figure 1.1. Experiment design.** The top panel (a) shows the study procedure for all three experiments, including an example encoding sequence (left) and memory test trials (middle & right). During the encoding phase, in Experiment 1 & 3, participants had to rely on memory when making responses during predictable events, but were provided with cues during random events. In Experiment 2, all aisle responses were cued. Panel (b) illustrates the relationship between item and action sequences. The learned action sequence is hypothesized to provide a scaffold for novel items seen throughout a single errand/event.

### Stimuli

Stimuli consisted of colorful, high-resolution images of objects from two semantic categories: animals (e.g., dog, elephant) and foods (e.g., banana, cupcake). There were 192 unique object images (96

per category) in total. All images had their backgrounds removed via Adobe Photoshop CC (2021 release). For the pretraining phase (see below), we also collected 24 abstract fractal images.

### **Task design**

**Overview.** Participants embarked on a series of errands to two different stores: the pet store and the grocery store. On each errand (i.e., event), their task was to “collect” a sequence of items by visiting a sequence of aisles (via keyboard button presses) within the current store. Importantly, events in one of these stores were defined as *predictable*, in that each time the participant ran an errand in this store, the sequence of aisles they visited followed a fixed, previously-learned order. Within the *random* store, however, participants visited a novel sequence of aisles during every event/errand they experienced. The identity of the predictable store (pet vs. grocery store) was counterbalanced across participants.

**Pretraining.** Before beginning their errands, participants completed a pretraining task, where they learned the aisle/action sequence they would follow in the predictable store. The pretraining phase consisted of three cycles of alternating study and test periods, during which participants repeatedly executed the sequence of actions (i.e., keyboard presses) associated with the predictable store. Each study repetition began by displaying a banner for 3s with the name of the store that was about to be visited, before moving on to a screen that displayed the four aisles within that store as four side-by-side circles (Figure 1.1a). Participants were then cued to visit a sequence of six aisles one at a time. Specifically, prior to each response, the circle associated with the to-be-visited aisle turned red, cueing participants to “enter” that aisle by pressing the corresponding key (d, f, j, or k). Only correct responses were accepted, and participants were encouraged to respond as quickly as possible. Trials were self-paced (up to 6s per aisle response), and a period of 500ms separated each response from the next aisle cue. Participants visited six aisles sequentially in this manner before moving on to the next study repetition. Throughout each repetition, a static, grayscale image of the store interior (pet or grocery store) remained on the screen in the background, providing a constant reminder of the current store’s identity.

After finishing a set of study repetitions, participants completed a 20s distractor task (see below). Next, during each pretraining test period, they had to recreate the aisle sequence (i.e., the sequence of

keyboard button presses) that was associated with the predictable store from memory. As in the study period, each test trial began with a banner depicting the name of the to-be-tested store for 3s. Then, participants had to press the sequence of buttons associated with that store's aisle sequence. Visual feedback appeared on the screen for 500ms after each response (green checkmark for correct responses; red X for incorrect responses). Participants recreated the predictable store's aisle sequence twice in this manner before receiving feedback about their accuracy across both test repetitions (out of 100%).

The entire pretraining phase consisted of three study-test cycles/blocks (Figure 1.2a). In the first two study periods (*pred-only* blocks), participants visited the predictable store five times in a row, and then were tested on the corresponding aisle sequence twice. However, during the third study period — the *intermixed* block — participants visited the predictable and random stores two times each (order randomized). This intermixed block also introduced a second new feature: after each aisle visit, participants viewed an abstract fractal image “in” that aisle (i.e., above the aisle circle) for 2.5s before the next aisle cue appeared. Participants were explicitly told to observe these abstract images without trying to remember them. These novel additions to the intermixed block’s study period — the inclusion of random store visits and of abstract stimuli — served two purposes, respectively: (1) to ensure that participants’ memory for the predictable sequence was robust to interference (from executing random sequences); and (2) to introduce the structure of the subsequent encoding task (described below), in which participants would see (i.e., “collect”) a novel item after each aisle visit. Both pred-only and intermixed study periods were followed by the same distractor task and test period, as described above. Across the entire pretraining period, therefore, participants completed 12 study repetitions and 6 test repetitions of the predictable aisle sequence.

**Encoding.** After the pretraining phase, the errand-running task began (Figure 1.1). As during pretraining study periods, participants first saw a banner for 3s depicting the upcoming store’s identity, before then visiting a sequence of six aisles by pressing buttons on their keyboards. Critically, however, this procedure differed for predictable versus random events. During random events, participants continued to see cues (i.e., red circles) specifying which aisle should be visited next. But in the

predictable store, they instead had to execute the associated aisle sequence *from memory*. In this case, rather than a red circle cue, a red fixation cross appeared in the center of the screen to mark the onset of the aisle response period. Responses in both conditions were self-paced, with a maximum time limit of 6s. All aisle responses were followed by trial-by-trial feedback similar to what was provided during the pretraining test periods (i.e., a green checkmark/red X for a correct/incorrect response). This feedback remained on the screen for 300ms, followed by an additional 200ms inter-trial-interval (ITI).

After making each aisle response, participants went on to “collect” a novel item from the store by viewing it in the given aisle (i.e., above the aisle circle); Figure 1.1a. Each item remained on the screen for 2.5s, and a 500ms ITI separated the end of the item presentation period and the start of the next trial. Items belonged to the semantic category consistent with the store’s identity: foods in the grocery store and animals in the pet store. A static image of the store interior also remained in the background during each event. Crucially, *all items presented during these events were novel*, irrespective of condition (predictable vs. random). Again, each event involved six aisle visits and six novel items. Prior to encoding, participants were instructed to remember both the *order* of the stimuli they collected, and the *aisles* that each item had been collected from. They were also told to imagine interacting with each item as it appeared on the screen, in order to encourage deeper encoding.

Participants completed a total of eight errand blocks throughout the experiment. Each encoding block included four errands/events: two in the predictable store, and two in the random store. The order of events was counterbalanced across blocks and participants, as was the order of object stimuli. After each encoding block, participants engaged in a short distractor task and then completed a set of memory tests for the items collected in that block (both tasks described below).

**Memory tests.** After each encoding block, participants completed two memory tests about the items collected during their recent set of errands: an order reconstruction test and a spatial memory test (Figure 1.1a). During the order reconstruction test, participants viewed all 6 of the collected items from a single event on the screen, and were told to reconstruct their temporal order by clicking on them in the order that they had been presented during encoding. A text prompt appeared on the upper half of the

screen to clarify which item in the sequence should be selected next (e.g., “click the *first* item you saw”). The order/position of items as they appeared on the screen was randomized, and participants had a maximum of 15s to make each response. One predictable event and one random event (out of four events total per block) were included in each block of the reconstruction test. During the spatial memory test, we probed participants’ memory for the aisles (i.e., spatial locations) that each item had been collected from. On each trial, a single item appeared in the center of the screen, and participants had a maximum of 8s to indicate via keyboard button press which of the four aisles it was associated with. Only items from events that were *not* included in the order reconstruction test (one predictable and one random event per block) appeared in this spatial memory test, in order to avoid overlap across the two memory judgments.

**Distractor task.** Participants completed a brief distractor task between study and test periods of the pretraining phase, and between encoding and retrieval blocks of the main errand task. This task was intended to disrupt any influence of recency effects and/or rehearsal on performance during the pretraining test periods or the memory tests. On each trial, a random single-digit number was presented in the center of the screen for 1250ms (with 250ms blank ITIs). Participants were told to press the space bar every time an even number appeared. Each repetition of the distractor task lasted approximately 20s.

**Reminder task.** Although participants saw trial-by-trial feedback during encoding regarding the accuracy of their aisle responses, we also included a brief reminder task after the fourth retrieval block (halfway through the experiment) to refresh memory for the predictable aisle sequence (Figure 1.1a). During this task, participants first passively observed the predictable sequence (as a series of red circle cues) twice in a row. Both of these sequence presentations were preceded by the appearance of a banner for 3s depicting the predictable store’s identity. Each aisle cue appeared on the screen for 1s, followed by a 500ms ITI. No responses were collected during this period, nor were any stimuli presented after each aisle cue. Afterward, participants completed two test trials in which they had to recreate the aisle sequence twice from memory. These test trials followed the same structure and timing of the pretraining test periods, with the exception that no information about overall accuracy was provided at the end of the reminder task. Results from the reminder task are discussed in Appendix A.

**Aisle sequence generation.** As described, each store contained 4 aisles and each errand was composed of 6 sequential aisle visits (i.e., actions). Aisle/action sequences were generated with the following constraints: (1) 2 out of 4 aisles were visited twice per sequence; (2) the same aisle could not be visited twice consecutively; (3) sub-sequences that traversed linearly through 3+ adjacent aisles were forbidden (e.g., a sub-sequence that moved from the leftmost aisle immediately to the middle-left aisle and then to the middle-right aisle); and (4) the first aisle visited in each sequence was never revisited (during the same event). From within the pool of valid sequences, two were randomly selected to serve as predictable aisle sequences: one beginning with an aisle on the edge of the screen, and one beginning with an aisle in the middle. Each participant was randomly assigned one of these two predictable sequences.

A set of random aisle sequences were then selected for each participant from the same pool, but with the following additional limitations: (1) random sequences could not begin with the same first aisle as the given participant's predictable sequence; (2) random sequences could not contain more than 3 aisles with the same ordinal position as the predictable sequence; and (3) random sequences could not share more than 4 aisles in the same ordinal position with each other. These additional rules helped to ensure not only that the predictable sequence was as distinct as possible from all of the random sequences seen by a given participant, but also that random sequences were dissimilar enough to each other to prevent the unintended learning of any pattern across them. A total of 18 random sequences were selected for each participant (2 of which appeared during pretraining, and 16 of which appeared during encoding).

### ***Statistical analysis***

For our main analyses, paired sample t-tests (two-tailed) were used to assess whether behavior differed as a function of event condition (i.e., predictable vs. random). Each of these t-tests indicates the corresponding 95% confidence interval (*CI*). However, when such analyses involved reaction time (RT) data, we instead used non-parametric, Wilcoxon signed-rank tests to account for the fact that RT distributions are typically skewed, and thus often violate the assumptions of normality made by parametric tests. Repeated-measures ANOVAs were also used to examine how accuracy differed by aisle location. For effect sizes, we report Cohen's *d* for t-tests (Cohen, 1988), Cliff's delta for Wilcoxon

signed-rank tests (Cliff, 1996), and partial eta-squared for ANOVAs (Cohen, 1973), along with corresponding 95% confidence intervals for each metric ( $CI_d$  for Cohen's d and Cliff's delta,  $CI_\eta$  for partial eta-squared). In the case of negative Cohen's d or Cliff's delta values, the sign was flipped such that reported effect sizes are always positive.

Where applicable, p-values were adjusted for multiple comparisons by controlling for the false discovery rate (FDR) (Benjamini & Hochberg, 1995). Corrected confidence intervals for these comparisons were computed by following the procedure in Benjamini et al. (2005) for calculating false coverage-statement rate (FCR) adjusted CIs. Specifically, given the number of comparisons ( $m$ ), the number of effects that reached significance after FDR correction ( $R$ ), and a significance level ( $q$ ), we report corrected CIs ( $CI_{FCR}$ ) with  $width = [100 * (1 - R * q/m)]\%$ . All statistical analyses were conducted in R (ver. 4.1.1; R Core Team, 2021) using functions from R's built-in *stats* package (ver. 4.1.1), *rstatix* (ver. 0.7.0; Kassambara, 2021), *effsize* (ver. 0.8.1; Torchiano, 2020), *Rmisc* (ver. 1.5.1; Hope, 2022), *apaTables* (ver. 2.0.8; Stanley, 2022), *lme4* (see below), and *brms* (see Appendix A).

**Multilevel logistic regression.** We also ran a multilevel regression model to explore how order memory performance varied as a function of both condition (predictable vs. random) and sequence position (1-6). In this model, binary accuracy on each memory test trial was predicted by condition (effect-coded), sequence position (mean-centered), and their interaction, with a random intercept for each subject, as well as both fixed and random slopes for each predictor. To allow for model convergence, we did not estimate correlations between random effects. This model was implemented using the *glmer* function from the *lme4* package (ver. 1.1-27.1, Bates et al., 2015). For each predictor, we report the associated beta coefficient ( $b$ ), standard error ( $SE$ ), 95% profile likelihood confidence interval ( $CI$ , as implemented by the *confint* function), and  $p$  value (estimated based on asymptotic Wald tests, as implemented by the *glmer* function). Note that alternate methods of computing  $p$  values for multilevel model terms (e.g., likelihood ratio tests) produced nearly identical results. We also used Bayesian multilevel regression to examine the effect of predictable action sequences on order memory while controlling for a number of different confounds (see Appendix A for details).

### **Data inclusion**

All data inclusion criteria were determined *a priori* and were consistent across Experiments 1, 2, and 3. For all pretraining analyses, we included all participants who made at least 75% of aisle responses during the study portion of the pretraining task ( $N = 80$ ). For all encoding analyses, we included all participants who (1) reached at least 80% accuracy on the final pretraining test repetition, and (2) made at least 75% of aisle responses during encoding ( $N = 69$ ). For analysis of each memory test, we included all participants who (1) reached at least 80% accuracy on the final pretraining test repetition, (2) made aisle responses in the predictable store with at least 80% accuracy, and (3) performed with above-chance accuracy on that test, according to a binomial test ( $N = 63$  in both the order reconstruction and spatial memory tests). Finally, for reminder task analyses (see Appendix A), we included the same participants as those in the memory test analyses, prior to any test-specific exclusions ( $N = 66$ ). Collectively, these criteria ensured not only that our sample of online participants followed task instructions (e.g., by making responses when required), but also that they acquired and retained memory for predictable aisle/action sequence throughout the experiment — a necessary prerequisite to test our hypotheses of interest.

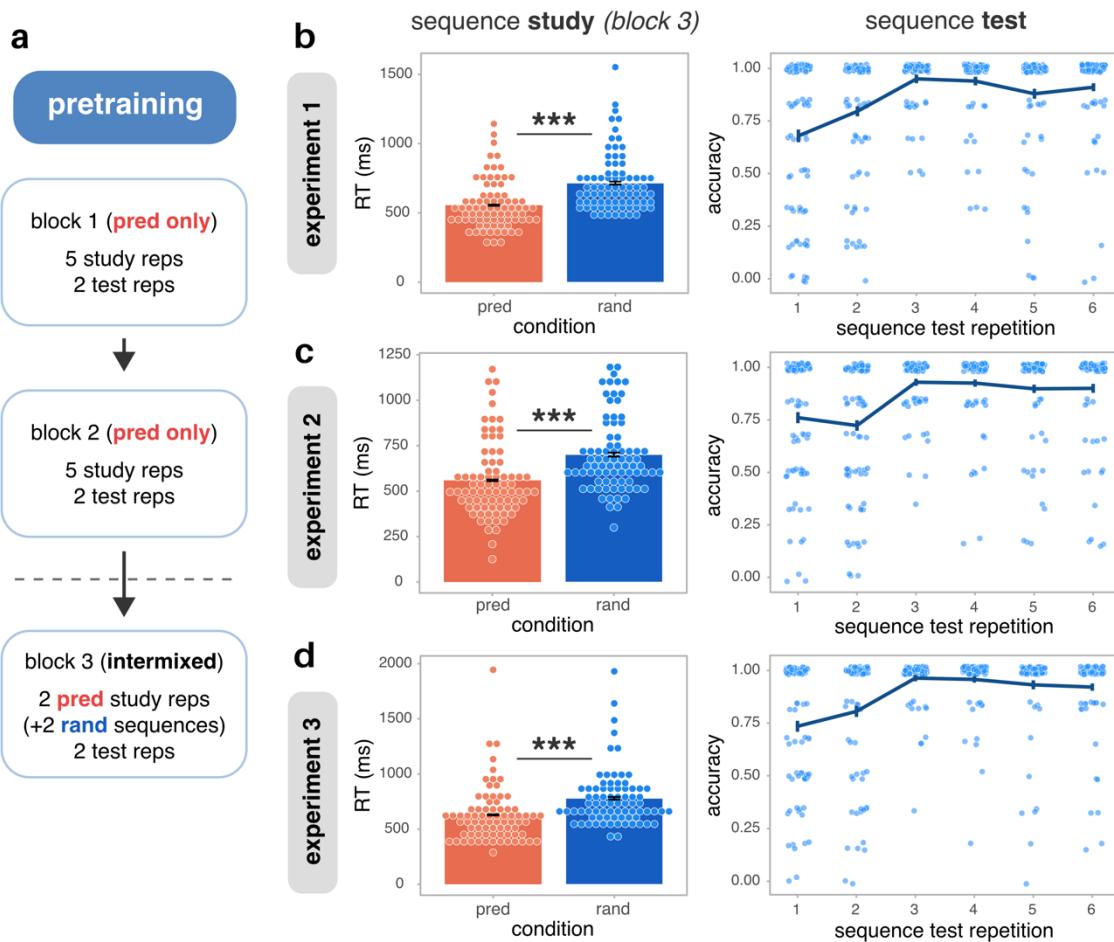
We also implemented the following trial-level exclusions. For all analyses involving RTs, outlier data points (defined as  $> 3$  SDs above or below the mean) were removed. When examining data from the memory tests (e.g., order reconstruction), we also excluded all trials with RTs lower than 100 ms, as these responses were implausibly fast and thus likely to be made in error. Such responses were exceedingly rare ( $< 0.9\%$  of trials), and did not change our findings when included.

## **1.2 Results**

### ***Pretraining behavior***

Prior to encoding, participants learned the action sequence associated with the predictable store during a pretraining phase, which consisted of three study-test blocks (Figure 1.2a). We assessed learning in two ways. First, in the final study block, we examined response times (RTs) as a function of store condition (predictable or random). RTs were faster for predictable versus random trials on average (Wilcoxon signed-rank test:  $V = 160$ ,  $N = 80$ ,  $p < 0.001$ , Cliff's  $d = 0.52$ ;  $CI_d = [0.35, 0.65]$ , Figure 1.2b)

and within each sequence position (Appendix A, Figure 1a), providing evidence of learning. Second, during the final test block, we considered the accuracy with which participants reproduced the learned sequence from memory. Participants were highly successful during this test, with mean accuracy reaching 0.91 ( $SD = 0.21$ ), and 88% ( $N = 70$ ) of our sample performing with at least 80% accuracy (Figure 1.2b). Importantly, only participants who demonstrated strong memory for the predictable aisle sequence by the end of pretraining (and who continued to execute that sequence accurately during encoding) were considered in subsequent analyses (see *Methods* for details).



**Figure 1.2. Pretraining design and behavior.** The left panel **(a)** shows the pretraining procedure for all experiments. Participants completed three cycles of study and test periods to allow for robust learning of the predictable store's aisle sequence. Pretraining performance is plotted separately for Experiment 1 **(b)**, 2 **(c)**, and 3 **(d)**. Leftmost plots show mean reaction times (RT) for predictable versus random aisle responses made during the study period of the third (“intermixed”) block of pretraining. Rightmost plots show accuracy across sequence test repetitions. Error bars indicate within-subject standard errors, and dots represent individual participants. \*\*\*  $p < 0.001$ .

### ***Learning of action sequences persists throughout encoding***

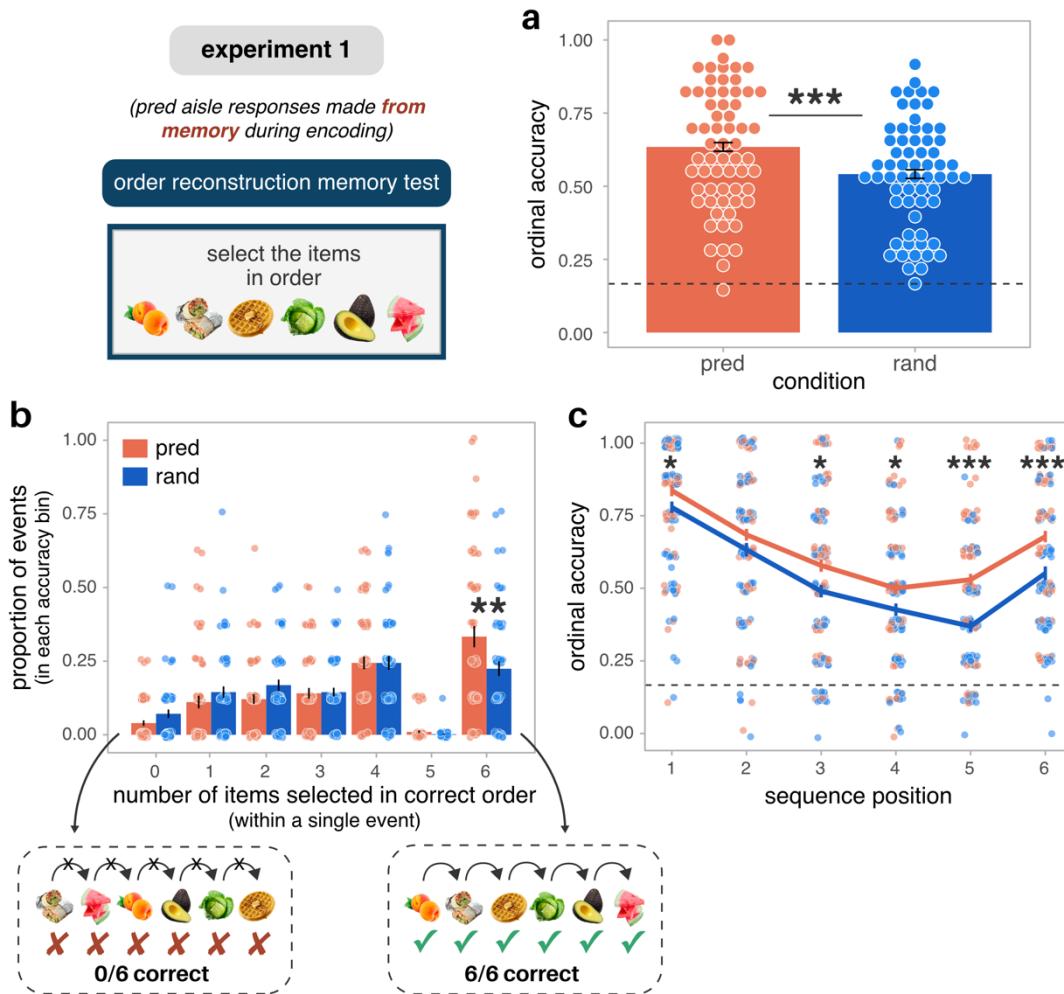
We also examined accuracy and RTs during the main errand-running task for both predictable and random aisle sequences. Accuracy was near ceiling in both conditions (predictable:  $M = 0.96$ ,  $SD = 0.08$ ; random:  $M = 0.98$ ,  $SD = 0.02$ ). Although participants were more accurate on average when making random versus predictable responses ( $t(68) = -2.59$ ,  $p = 0.012$ ,  $CI = [-0.04, -0.01]$ ,  $d = 0.36$ ,  $CI_d = [0.08, 0.64]$ ) — a finding that is not unsurprising given that random responses were explicitly cued — this accuracy difference disappeared rapidly as participants gained experience with the task (Appendix A, Figure 2a). Namely, responses during random events were significantly more accurate than those during predictable events only in the first block ( $t(68) = -4.13$ ;  $p_{FDR} < 0.001$ ,  $CI_{FCR} = [-0.18, -0.03]$ ,  $d = 0.72$ ,  $CI_d = [0.08, 0.64]$ ; all other blocks:  $|t(68)| < 2$ ;  $p_{FDR} > 0.2$ ).

Throughout the majority of the encoding task, therefore, participants were just as accurate when making aisle responses from memory as they were when responding to cues. Nevertheless, it remains possible that responses in the random/cued condition may still have been easier for participants to make. To address this possibility, we examined aisle response RTs for predictable versus random trials. No difference in RTs was observed (Wilcoxon signed-rank test:  $V = 1112$ ,  $N = 69$ ,  $p = 0.57$ , Cliff's  $d = 0.04$ ,  $CI_d = [-0.15, 0.22]$ ), indicating that although predictable and random encoding trials differed in their task demands, participants' actions within each condition were comparable in both accuracy and speed.

### ***Predictable action sequences scaffold temporal order memory***

After each encoding block, participants completed an order reconstruction test (Figure 1.3). Our critical question is whether engagement in a known (motor) action sequence during encoding leads to a cross-modal enhancement in temporal memory for novel and unrelated (visual) items. We first looked at ordinal accuracy, or the proportion of trials during which participants selected the correct item in the correct ordinal position (e.g., first, second). We found that ordinal accuracy was significantly higher for predictable than for random events,  $t(62) = 4.45$ ,  $p < 0.001$ ,  $CI = [0.05, 0.13]$ ,  $d = 0.47$ ,  $CI_d = [0.25, 0.69]$  (Figure 1.3a). This enhancement in temporal memory for predictable events was also observed when using Levenshtein distance (Levenshtein, 1966) as our index of order accuracy, ( $t(62) = -4.07$ ,  $p < 0.001$ ,

$CI = [-0.65, -0.23]$ ,  $d = 0.43$ ,  $CI_d = [0.21, 0.65]$ ). Levenshtein distance refers to the minimum number of “edits” (insertions, deletions, or subtractions) between the actual and reconstructed sequences, and more readily accounts for situations in which participants correctly recalled sub-sequences of items but selected them in the wrong ordinal positions.



**Figure 1.3. Order reconstruction memory results for Experiment 1.** The bar plot in (a) shows ordinal accuracy for novel items from predictable versus random events. The bar plot in (b) shows the proportion of events in which  $n/6$  items ( $n = 0-6$ ) were selected in the correct ordinal position. (Note that the near-zero proportion of events in which 5 items were remembered stems from the fact that getting exactly 5 out of 6 responses correct required either missing one response or selecting the same item twice in different ordinal positions, both of which rarely occurred.) The plot in panel (c) shows ordinal accuracy as a function of sequence position within the event. Significance values are corrected for multiple comparisons. Error bars indicate within-subject standard errors, and dots represent individual participants. Dotted lines indicate chance performance ( $1/6 = 0.17$ ). \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Importantly, the effect of condition on order memory held after controlling for numerous confounds (sequence position, event position, block number, and item category), and when considering memory only for items that were preceded by accurate aisle responses during encoding (Appendix A, Table 1). We also sought to account for the fact that during encoding, each 6-item aisle sequence included two aisles that were visited once, and two that were visited twice. This aisle repetition structure was present in both predictable and random sequences, making it unlikely to explain our order memory effects. Nevertheless, it is possible that the association of two novel items with the same spatial location produced interference in memory, which could have different effects within each condition. To rule out this possibility, we examined ordinal accuracy as a function of whether items were associated with aisles that were visited once versus twice per event. Although ordinal accuracy was generally lower for items associated with repeated aisles, suggesting that aisle repetition did indeed generate interference, the predictable versus random order memory benefit was robust to this confound (Appendix A). Taken together, these results provide strong evidence that the execution of a familiar action sequence scaffolds temporal memory for simultaneously-encoded novel items.

#### ***Predictable action sequences promote the formation of holistic event memories***

According to our scaffolding hypothesis, engaging in a familiar sequence of actions during encoding facilitates temporal order memory for unrelated visual items by allowing them to be integrated within a pre-existing event memory, where each novel item gets slotted into a position of the action sequence representation (Figure 1.1b). This framework predicts that sequences from predictable events are more likely to be remembered in their entirety, given that each novel visual item from the sequence can be integrated into the same, stable event representation. Indeed, the ability to recall the complete set of elements from an event is a hallmark of intact episodic memory (Horner et al., 2015). To test this prediction, we examined order accuracy at the level of entire events. For each participant, we determined the proportion of events in which  $n/6$  items (where  $n = 0$  through 6) were selected in the correct ordinal position (Figure 1.3b). We found that the proportion of events that were reconstructed in their entirety was greater in the predictable versus random condition ( $t(62) = 3.57$ ,  $CI_{FCR} = [0.02, 0.19]$ ,  $p_{FDR} = 0.005$ ,  $d$

$= 0.44$ ,  $CI_d = [0.19, 0.70]$ ). In contrast, there were no condition-related differences in the proportion of events from which participants accurately reported the order of 0-5 items (all  $|t(62)| < 2.3$ ,  $p_{FDR} > 0.05$ ).

To explore whether any position effects were evident, we also looked at order accuracy separately for each sequence position (1-6). Order memory for items in all positions, except the second, was higher for predictable versus random events (position 1:  $t(62) = 2.13$ ,  $p_{FDR} = 0.045$ ,  $CI_{FCR} = [0.001, 0.11]$ ,  $d = 0.27$ ,  $CI_d = [0.01, 0.53]$ ; position 2:  $t(62) = 1.60$ ,  $p_{FDR} = 0.11$ ,  $CI_{FCR} = [-0.01, 0.12]$ ,  $d = 0.20$ ,  $CI_d = [-0.05, 0.45]$ ; position 3:  $t(62) = 2.68$ ,  $p_{FDR} = 0.019$ ,  $CI_{FCR} = [0.02, 0.16]$ ,  $d = 0.35$ ,  $CI_d = [0.08, 0.61]$ ; position 4:  $t(62) = 2.24$ ,  $p_{FDR} = 0.043$ ,  $CI_{FCR} = [0.01, 0.14]$ ,  $d = 0.30$ ,  $CI_d = [0.03, 0.57]$ ; position 5:  $t(62) = 5.33$ ,  $p_{FDR} < 0.001$ ,  $CI_{FCR} = [0.10, 0.22]$ ,  $d = 0.70$ ,  $CI_d = [0.41, 0.98]$ ; position 6:  $t(62) = 4.12$ ,  $p_{FDR} < 0.001$ ,  $CI_{FCR} = [0.06, 0.19]$ ,  $d = 0.53$ ,  $CI_d = [0.26, 0.80]$ ), suggesting that action sequences scaffold memory throughout the novel item sequence (Figure 1.3c).

However, qualitative examination of the data in Figure 1.3c suggests that participants' temporal memory benefited from the predictable action sequence to a larger degree for items toward the *end* of the six-item event. To quantify this interactive effect, we ran a multilevel logistic regression model, in which ordinal accuracy (on each trial) was predicted by condition, sequence position, and their interaction (with subject-specific random effects for each predictor and a random intercept for each subject). This model revealed strong main effects of both condition ( $b_{cond} = 0.49$ ,  $SE = 0.11$ ,  $CI = [0.28, 0.71]$ ,  $p < 0.001$ ) and sequence position ( $b_{seq\_pos} = -0.24$ ,  $SE = 0.02$ ,  $CI = [-0.28, -0.20]$ ,  $p < 0.001$ ) on memory performance, such that participants were more accurate for items from predictable events *and* from earlier sequence positions. Critically, the interaction between these variables was also significant ( $b_{int} = 0.07$ ,  $SE = 0.03$ ,  $CI = [0.003, 0.14]$ ,  $p = 0.038$ ), confirming that the predictable  $>$  random order memory effect was indeed stronger at the end of a sequential event compared to the beginning. Such results can also be interpreted in light of the scaffolding hypothesis. As participants progress deeper into an event, the passage of time is likely to cause a gradual drift in their mental context; larger contextual drifts within an event, in turn, may disrupt temporal binding (Clewett & Davachi, 2017; DuBrow et al., 2017; Howard & Kahana, 2002). By anchoring participants to a familiar and well-defined event representation, it could be that the predictable

action sequence mitigates the extent of this contextual drift, boosting the likelihood that items toward the end of a sequence can be effectively bound to the full event.

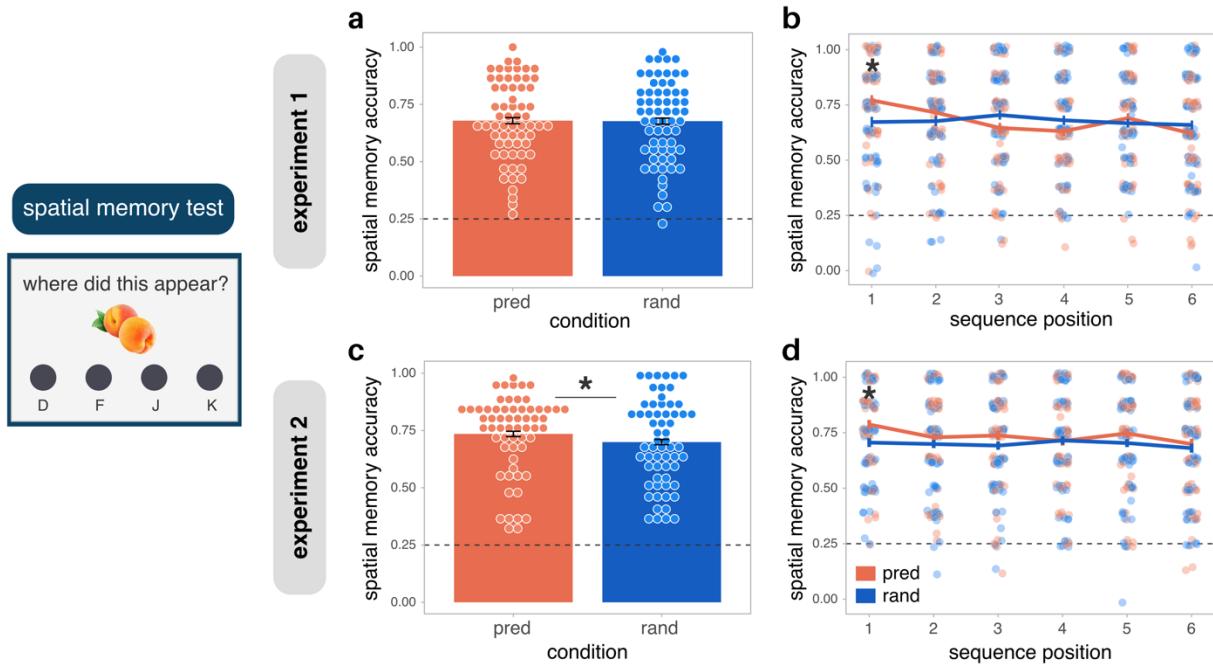
### ***Effects of action sequences on spatial memory***

Thus far, we have established that when participants encounter novel information in the context of a known action sequence, they are better able to remember the temporal links between those items. However, we saw no similar benefit in spatial context memory. Spatial memory performance was comparable for predictable versus random events when averaging across items from all sequence positions,  $t(62) = 0.11, p = 0.91, CI = [-0.04, 0.04], d = 0.01, CI_d = [-0.04, 0.04]$  (Figure 1.4a). Breaking down memory performance by sequence position, we do find evidence that spatial memory for items at the start of each errand was better for predictable versus random events (position 1:  $t(62) = 2.95, p_{FDR} = 0.027, CI_{FCR} = [0.01, 0.19], d = 0.41, CI_d = [0.12, 0.69]$ ); all other positions:  $|t(62)| < 2, p_{FDR} > 0.1$ ; Figure 1.4b. Although this effect is intriguing, we do not interpret it too strongly because it is partly confounded by the fact that participants' spatial memory performance also differed as a function of aisle location (with memory for items in the left-most aisle being better than other aisles on average), and that the location of the first aisle in each sequence was not matched across predictable and random events (Appendix A). However, future work could explore how the onset of a familiar action sequence might trigger enhanced spatial context memory.

### **1.2.3 Discussion**

Experiment 1 demonstrates that executing a well-learned action sequence significantly improves memory for the temporal order of novel, non-motor items encountered simultaneously — but does not improve memory for items' spatial position. Next, we asked whether this order memory effect was dependent on the need to *use* one's memory for the familiar action sequence during encoding. In Experiment 1, participants made aisle responses in the predictable store *from memory*, whereas responses in the random store were cued. Although this difference in task demands did not induce condition-related differences in the speed or accuracy of participants' actions, it raises an intriguing question: will the execution of a familiar action sequence scaffold temporal memory for novel items even when knowledge

of that sequence is not necessary for behavior? To this end, in Experiment 2 we again trained participants on the predictable action sequence, but then eliminated the need to retrieve aisle responses from memory by providing cues in both conditions.



**Figure 1.4. Spatial memory test performance.** Data from Experiments 1 (top) and 2 (bottom). Panels (a) and (c) show mean accuracy as a function of store condition, while (b) and (d) show accuracy in each condition as a function of the sequence position in which an item was encountered during encoding (with significance values FDR-corrected for multiple comparisons). Error bars indicate within-subject standard errors, and dots represent individual participants. Dotted lines indicate chance performance ( $1/4 = 0.25$ ). \*  $p < 0.05$ .

## 1.3 Experiment 2

### 1.3.1 Methods

#### *Participants*

Participants were recruited again through Prolific ( $N = 80$  participants; 32 female). All participants were healthy adults between the ages of 18-35 ( $M = 26.4$ ,  $SD = 5.0$ ), reported having normal or corrected-to-normal vision, and had a Prolific study approval rate of at least 60% prior to enrolling. Participants received \$14 upon study completion (completion time  $M = 61.6$  min,  $SD = 23.6$ ). All participants provided informed consent, and all procedures were approved by the Institutional Review

Board at Columbia University.

### **Task design**

Experiment 2 followed the same procedure as Experiment 1, with the following modifications.

First, both predictable and random events included the presentation of aisle cues that instructed participants which aisle to visit on every trial, thus removing the need to use memory for the learned action sequences during the critical errand encoding phase. We also removed trial-by-trial feedback, given that such feedback should be unnecessary when explicit cues are available to inform participants of all correct responses (and given that accuracy in the cued/random condition in Experiment 1 was essentially at ceiling). To further ensure that participants executed the correct sequence of actions within each event, even without feedback, only correct aisle responses were accepted and served to advance the trial to the item collection/presentation phase. In addition to these changes to the encoding phase, we also removed the reminder task after the end of the fourth block, given that all predictable events provided participants with repeated opportunities to observe and execute the predictable aisle sequence. We also added a “final” order reconstruction memory test at the very end of the experiment, in which we re-assessed order memory for half of the events that participants had already been tested on. This final test is not detailed in the current paper, but we note that order memory performance at this timepoint was enhanced for predictable versus random events, consistent with results from the order reconstruction tests that followed each errand block (see *Results*).

### **Data inclusion**

Experiment 2 applied identical data inclusion criteria as Experiment 1. This resulted in  $N = 80$  participants included in all pretraining analyses,  $N = 67$  in all encoding analyses, and  $N = 61$  in both order reconstruction test and spatial memory test analyses. Trials during pretraining or encoding with aisle response RTs greater than 3 SDs above the mean were removed as outliers from all RT analyses. For all memory test analyses, trials with implausibly fast responses were also removed ( $< 0.4\%$  of trials; again, keeping these trials in our analyses did not alter our results).

### 1.3.2 Results

#### ***Pretraining and encoding behavior***

As in Experiment 1, participants in Experiment 2 effectively learned the predictable aisle sequence by the end of pretraining. During the final study period, in which participants were cued to visit aisles in both predictable and random stores (see Figure 1.2a), RTs were consistently faster for predictable visits on average ( $V = 207, N = 80, p < 0.001$ , Cliff's  $d = 0.46, CI_d = [0.28, 0.60]$ ; Figure 1.2c) and within each sequence position (Appendix A, Figure 1b). During the final pretraining test trial, performance was high ( $M = 0.90; SD = 0.22$ ) with 84% ( $N = 67$ ) of participants recalling the sequence with at least 80% accuracy. Like in Experiment 1, only participants who successfully learned the predictable aisle sequence by the end of pretraining were considered in subsequent analyses (see *Methods* for details).

We next turned to encoding periods to assess whether memory for this predictable action sequence persisted across errand blocks. In this experiment, participants' aisle responses were cued in both the predictable and random conditions, and only correct responses were accepted. Rather than examining response accuracy, therefore, we turned to the speed of aisle responses as a proxy for learning. Aisle response RTs were significantly faster during predictable versus random events (Wilcoxon signed-rank test:  $V = 235, N = 67, p < 0.001$ , Cliff's  $d = 0.21, CI_d = [0.02, 0.40]$ ). This effect reveals that memory for the predictable action sequence afforded predictive processes that allowed participants to deploy responses in this condition more rapidly — even though such predictions were not needed to guide behavior.

#### ***Predictable action sequences scaffold order memory even when not critical for behavior***

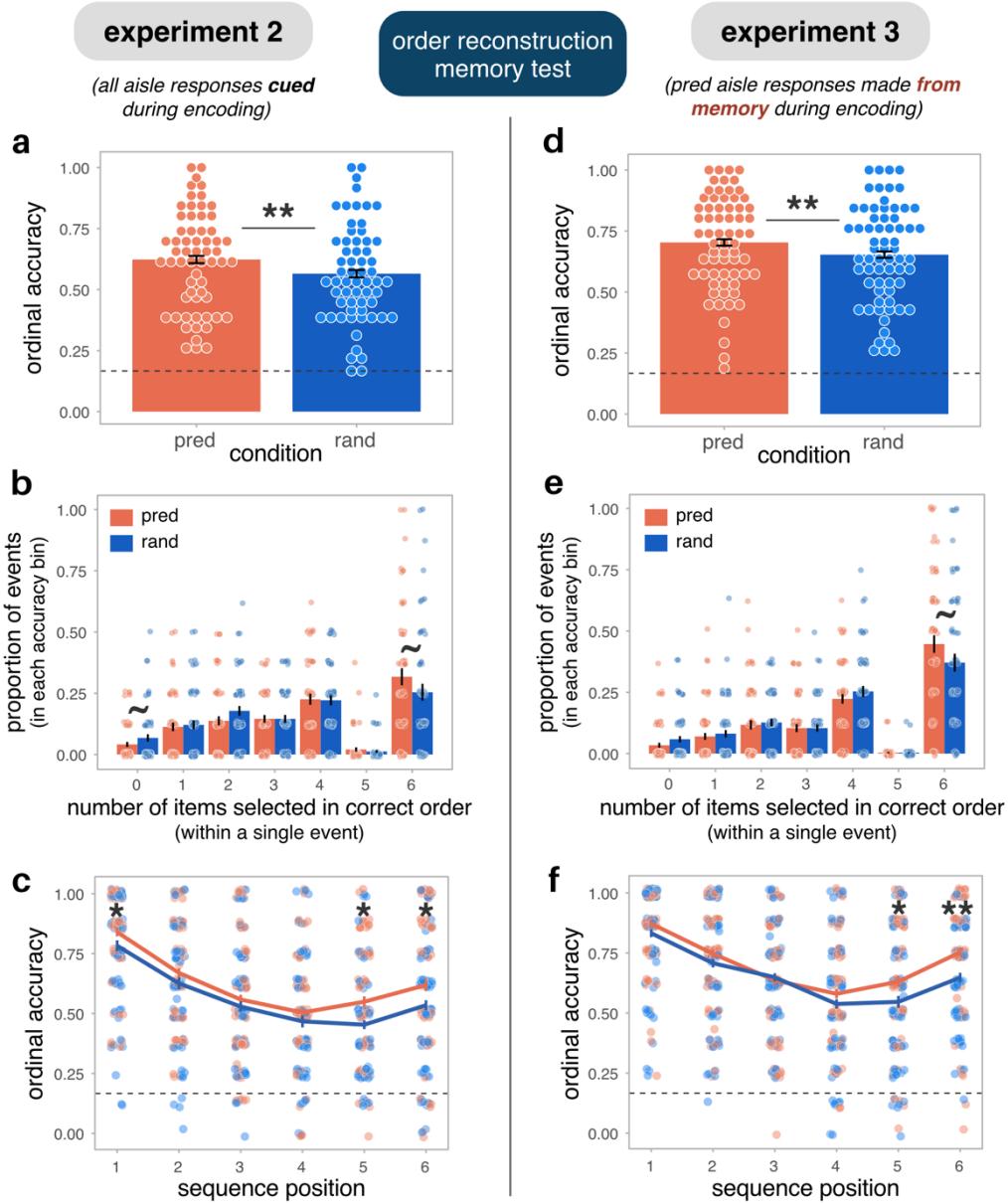
In this experiment, participants were not explicitly required to use their memory for the predictable action sequence during encoding. Nevertheless, here we replicate all of the main effects observed in Experiment 1 (Figure 1.5a-c). Performance on the order reconstruction test was enhanced for predictable versus random events, as revealed by both ordinal accuracy ( $t(60) = 2.74, p = 0.008, CI = [0.02, 0.10], d = 0.29, CI_d = [0.08, 0.50]$ ) and Levenshtein distance ( $t(60) = -2.80, p = 0.007, CI = [-0.54, -0.09], d = 0.29, CI_d = [0.08, 0.50]$ ). This effect was also robust to the influence of confound variables —

including whether items were associated with repeated versus non-repeated aisles (Appendix A, Table 2). These findings demonstrate that the mere presence of familiar sequential actions during novel encoding scaffolds memory for how events unfold across time, even if this prior knowledge is not necessary for an individual's moment-to-moment behavior.

As with Experiment 1, we also see evidence that participants were more likely to accurately reconstruct the complete sequence of novel items for predictable versus random events (Figure 1.5b), pointing to the capacity of a known action sequence to facilitate the formation of holistic event representations. Participants were again more likely to select 6/6 items in the correct order for predictable versus random events ( $t(60) = 2.15, p = 0.035, CI = [0.01, 0.12], d = 0.24, CI_d = [0.01, 0.47]$ ); however, this comparison did not survive FDR correction for multiple comparisons ( $p_{FDR} = 0.15$ ). There were also no significant condition-related differences in the proportion of events in which  $n = 0-5$  (out of 6) items were selected in the correct ordinal position (all  $|t(60)| < 2.1$ , all  $p_{FDR} > 0.15$ ).

Finally, we replicate the qualitative pattern that predictable action sequences might have a stronger temporal memory-enhancing effect for items toward the end of an event (Figure 1.5c). When examining accuracy separately for each sequence position, we find that only items in the first ( $t(60) = 2.41, p_{FDR} = 0.038, CI_{FCR} = [0.003, 0.11], d = 0.26, CI_d = [0.04, 0.47]$ ), fifth ( $t(60) = 3.27, p_{FDR} = 0.011, CI_{FCR} = [0.03, 0.16], d = 0.40, CI_d = [0.15, 0.65]$ ), and sixth ( $t(60) = 2.62, p_{FDR} = 0.033, CI_{FCR} = [0.01, 0.16], d = 0.30, CI_d = [0.07, 0.54]$ ) sequence position were significantly better remembered during predictable versus random events (all other positions:  $|t(60)| < 1.5, p_{FDR} > 0.2$ ). To better quantify the interactive effects of condition and sequence position, we again ran a trial-level logistic regression model with the same design as that used in Experiment 1. This model revealed significant effects of both condition ( $b_{cond} = 0.29, SE = 0.11, CI = [0.08, 0.50], p = 0.007$ ) and sequence position ( $b_{seq\_pos} = -0.24, SE = 0.02, CI = [-0.29, -0.19], p < 0.001$ ), such that participants showed a greater enhancement in order memory for predictable versus random items and for items at the end of an event. The interaction term in this model was again positive ( $b_{int} = 0.04, SE = 0.04, CI = [-0.03, 0.11], p = 0.29$ ) — indicating that predictable action sequences enhanced temporal memory more strongly for items toward the end of an

event — but this effect did not reach significance.



**Figure 1.5. Order reconstruction memory results for Experiments 2 and 3.** Order memory performance is plotted as mean ordinal accuracy for predictable versus random events, separately for Experiment 2 (**a**) and Experiment 3 (**d**). Plots in panels (**b**) and (**e**) show the proportion of tested events in which  $n/6$  items (where  $n = 0-6$ ) were selected in the correct ordinal position, for Experiment 2 and 3 respectively. Plots in panels (**c**) and (**f**) show memory performance within each condition as a function of sequence position (1-6). Significance values are corrected for multiple comparisons. Error bars indicate within-subject standard errors, and dots represent individual participants. Dotted lines indicate chance performance ( $1/6 = 0.17$ ).  $\sim p_{unc} < 0.05$ ,  $*$   $p < 0.05$ ,  $** p < 0.01$ .

When taken together, these results show that demand to use one's memory for a familiar action sequence is not a necessary prerequisite for its facilitatory effects on order memory. That is, even though participants received cues to guide their actions during predictable events, their pre-existing familiarity with this sequence nevertheless functioned as a scaffold for novel temporal event memories.

### ***Effects of action sequences on spatial memory***

Next, we turned to the spatial memory test. In Experiment 1, we saw no difference in average spatial context memory for predictable versus random events. However, in Experiment 2, spatial memory was slightly but significantly higher in the predictable condition,  $t(60) = 2.20, p = 0.032, CI = [0.003, 0.06], d = 0.20, CI_d = [0.02, 0.38]$  (Figure 1.4c). This effect was driven by enhanced spatial memory for the location of the first item within predictable events (as was also observed in the previous experiment, despite the lack of a significant difference when averaging across sequence positions),  $t(60) = 2.73, p_{FDR} = 0.050, CI_{FCR} = [0.0001, 0.16], d = 0.41, CI_d = [0.10, 0.71]$ ; Figure 1.4d. Performance did not differ across conditions in any other sequence position (all  $|t(60)| < 2, p_{FDR} > 0.2$ ). As discussed, this result is difficult to interpret in light of the fact that spatial memory also differed as a function of aisle location (see Appendix A), and that aisle locations were not balanced equally across sequence positions and conditions. The effect of condition on spatial memory in this experiment was also not significantly greater than the null effect in Experiment 1, further limiting our interpretation (Appendix A, Table 4).

### **1.3.3 Discussion**

Experiment 2 replicates Experiment 1, finding that familiar motor action sequences benefit memory for the order of novel, unrelated visual items. Further, we find that the demand to retrieve this action sequence from memory during encoding is not required to benefit temporal memory. These results suggest that familiar action sequence memories may be automatically or implicitly engaged during encoding, which in turn is sufficient to bolster temporal integration of novel items.

Thus far, our conclusions that familiar motor sequence execution selectively or disproportionately supports temporal order memory is consistent with the scaffolding hypothesis, which argues that predictable action sequences should specifically aid in the construction of temporally-coherent event

memories — as opposed to the attentional resource hypothesis, which would predict widespread memory benefits. However, spatial context memory performance does not directly reflect attention to individual items. Thus, to better adjudicate between our hypotheses, Experiment 3 tested detailed item memory in addition to order memory.

## 1.4 Experiment 3

### 1.4.1 Methods

#### *Participants*

We again used Prolific to recruit  $N = 80$  participants (47 female). All participants were healthy adults between the ages of 18-35 ( $M = 24.7$ ,  $SD = 4.8$ ), reported having normal or corrected-to-normal vision, and had a Prolific study approval rate of at least 60% prior to enrolling. Participants received \$14 upon study completion (completion time  $M = 57.5$  min,  $SD = 14.4$ ). All participants provided informed consent, and all procedures were approved by the Institutional Review Board at Columbia University.

#### *Method*

Experiment 3 followed largely the same procedure as Experiment 1, in that participants made aisle responses from memory when running errands in the predictable store (instead of seeing cues in both conditions, as in Experiment 2). Although both Experiments 1 and 2 revealed comparable order memory results (Appendix A, Table 4), effect sizes were numerically larger in the first — perhaps because the explicit demand to use action sequence knowledge during encoding increased its effectiveness as a memory scaffold. Given that Experiment 3 aimed to test whether familiar action sequences generally increase attention to/encoding of novel items, we chose the design in which we saw the most robust (temporal) memory effects. We also made two additional modifications to our procedure. First, the spatial memory test was removed, such that after each errand block, participants completed only the order reconstruction test. Second, we administered an item recognition test at the very end of the experiment.

#### *Item recognition test*

After completing all encoding and order reconstruction test blocks, participants first engaged in a

brief distractor task (lasting approximately 5 minutes). This task comprised a set of 20 simple addition and subtraction problems (e.g.,  $28 + 76$ ). Participants saw one math problem at a time and were given 15 seconds to type their response. Next, the item recognition test began. In this test, participants discriminated between previously-encoded items and similar lures. For this experiment, we thus created two sets of stimuli: set A and set B. These sets contained the same specific categories of stimuli (e.g., cat, burrito), but different exemplar images. Approximately half of participants saw stimulus set A during their errands, while matching lure stimuli were drawn from set B. The other half saw set B during their errands and had lures drawn from set A. During the test, participants viewed a series of individual items on the screen and indicated whether each one was OLD (i.e., had been seen/collected during one of their errands) or NEW (i.e., had never appeared throughout the experiment). Four response options allowed for varying levels of confidence: “definitely OLD”, “maybe OLD”, “maybe NEW”, and “definitely NEW” (Figure 1.6). Importantly, all NEW images belonged to the same “type” (e.g., cat, pizza) as one of the OLD images, and thus functioned as similar lures. See Figure 1.6c for examples. Thus, successful performance on this task required memory for the specific details of individual items, above and beyond memory for their verbal category labels or general semantics. Participants were explicitly told that NEW images might be very similar to items they had collected during their errands, but that they should only use the OLD response when an image was exactly the same as something they’d already seen.

The item recognition test included 96 trials: half contained items from the predictable store (or were the same item type as something seen in the predictable store), and the other half contained items associated with the random store. Within each condition, there were 24 OLD trials and 24 NEW trials (i.e., trials presenting similar lures). Both condition (predictable vs. random) and recognition trial type (OLD vs. NEW) were sampled evenly across encoding blocks and sequence positions (within each event/errand). Trials were separated by an ITI of 1s, and participants were given up to 8s to make each response. Importantly, there was no overlap between the item recognition and order reconstruction tests. That is, items from half of all events (one predictable and one random event per block) were included in the order reconstruction test, and the remaining half were reserved for the item recognition test.

### **Data inclusion**

Experiment 3 used the same data inclusion criteria as Experiments 1 and 2. This left  $N = 78$  participants in pretraining analyses,  $N = 70$  in encoding analyses,  $N = 67$  in reminder task analyses,  $N = 66$  in order reconstruction test analyses, and  $N = 53$  in recognition test analyses. As described in the *Methods* of Experiment 1, participants who did not perform with above-chance accuracy on a particular memory test, according to a binomial test, were removed from analyses of that test. The number who failed to meet this criterion for the recognition test ( $N = 13$ ) was notably larger than that observed for other memory tests (mean  $N = 3.8$  across tests and experiments), suggesting that item recognition was generally more difficult. However, the effects described in the *Results* do not change when including all low-performing participants in our recognition test analyses. Finally, trials during pretraining or encoding with aisle response RTs greater than 3 SDs above the mean were removed as outliers from all RT analyses, and trials with implausibly fast responses were removed from all memory test analyses (< 0.3% of trials; as with other experiments, keeping these trials in our analyses did not change our results).

### **1.4.2 Results**

#### ***Pretraining and encoding behavior***

Participants' behavior during pretraining replicated Experiments 1 and 2 and indicated effective learning of the predictable aisle sequence. Specifically, during the final pretraining study block, aisle response RTs were consistently faster for predictable versus random store visits, both on average ( $V = 226$ ,  $N = 78$ ,  $p < 0.001$ , Cliff's  $d = 0.44$ ,  $CI_d = [0.27, 0.59]$ ; Figure 1.2d) and across all sequence positions (Appendix A, Figure 1c). Accuracy during the final pretraining test was also very high ( $M = 0.92$ ,  $SD = 0.19$ ), with 91% ( $N = 71$ ) of participants reproducing the aisle sequence with at least 80% accuracy.

Experiment 3 followed largely the same procedure as Experiment 1, in that participants made aisle responses from memory when running errands in the predictable store (instead of seeing cues in both conditions, as in Experiment 2). Aisle response behavior during this experiment replicated the effects in Experiment 1. That is, aisle response accuracy in the cued, random condition was slightly greater than in the predictable condition ( $t(69) = -2.57$ ,  $p = 0.012$ ,  $CI = [-0.04, -0.004]$ ,  $d = 0.39$ ,  $CI_d = [0.08, 0.70]$ ), but

this difference disappeared after the first block of the experiment (block 1:  $t(69) = -3.41$ ;  $p_{FDR} = 0.009$ ,  $CI_{FCR} = [-0.18, -0.02]$ ,  $d = 0.54$ ,  $CI_d = [0.20, 0.87]$ ; all other blocks:  $|t(69)| < 2.4$ ;  $p_{FDR} > 0.07$ ); Appendix A, Figure 2b. We also found that participants did not significantly differ in their response times to make predictable versus random aisle responses (Wilcoxon signed-rank test:  $V = 1143$ ,  $N = 70$ ,  $p = 0.56$ , Cliff's  $d = 0.02$ ,  $CI_d = [-0.17, 0.21]$ ) — even though these conditions differed in their task demands. These two results demonstrate that participants were able to effectively execute the predictable action sequence from memory during encoding, and did so with similar accuracy and speed as they showed in the random, cued condition. As with previous experiments, only participants who demonstrated effective learning of the predictable aisle sequence during pretraining and maintained their memory throughout encoding were included in subsequent analyses (see *Methods*).

### ***Replication of temporal memory enhancement for predictable events***

The results from Experiment 3 replicate those observed in Experiments 1 and 2 (Figure 1.5d-f). Participants were again significantly more accurate in their reconstruction of the order of items from the predictable versus the random store, both according to ordinal accuracy ( $t(65) = -2.68$ ,  $p = 0.009$ ,  $CI = [0.01, 0.09]$ ,  $d = 0.25$ ,  $CI_d = [0.06, 0.44]$ ) and Levenshtein distance ( $t(65) = -2.58$ ,  $p = 0.012$ ,  $CI = [-0.43, -0.06]$ ,  $d = 0.24$ ,  $CI_d = [0.05, 0.43]$ ). This effect held after controlling for confound variables, when considering only trials where participants visited the correct aisle during encoding, and when looking separately at items associated with aisles visited once or twice per sequence (Appendix A, Table 3).

We also again replicate the finding that participants were more likely to remember the order of predictable events holistically (i.e., in their entirety) relative to random events (Figure 1.5e). That is, participants selected 6/6 items in the correct order more often when they had been encoded in the predictable versus random store ( $t(65) = 2.36$ ,  $p = 0.021$ ,  $CI = [0.01, 0.14]$ ,  $d = 0.27$ ,  $CI_d = [0.04, 0.50]$ ); although as with Experiment 2, this difference was only significant at an uncorrected threshold ( $p_{FDR} = 0.15$ ). There were no significant condition-related differences in the proportion of events in which n = 0-5 (out of 6) items were selected in the correct ordinal position (all  $|t(65)| < 2.4$ , all  $p_{FDR} > 0.14$ ). Taken together with previous experiments, we interpret this pattern of results to reflect that participants'

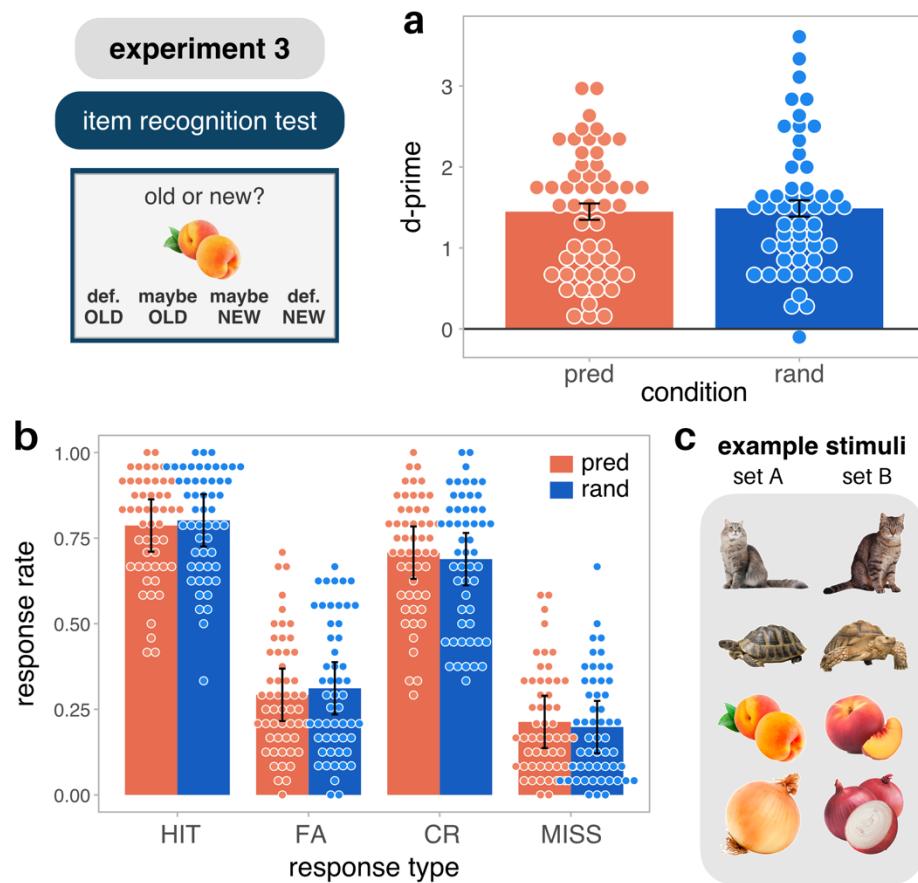
engagement in a familiar action sequence during a visual encoding event increases the likelihood that all elements of that event can be embedded within the same sequential memory representation.

When breaking ordinal accuracy down by sequence position, we also replicate the pattern that the order memory benefit for predictable versus random events was strongest for items toward the end of an event (Figure 1.5f). Specifically, the difference in ordinal accuracy across conditions was only significant for items in the fifth ( $t(65) = 2.36$ ,  $p_{FDR} = 0.021$ ,  $CI_{FCR} = [0.01, 0.18]$ ,  $d = 0.32$ ,  $CI_d = [0.04, 0.59]$ ) and sixth ( $t(65) = 3.71$ ,  $p_{FDR} = 0.004$ ,  $CI_{FCR} = [0.03, 0.18]$ ,  $d = 0.45$ ,  $CI_d = [0.20, 0.71]$ ) sequence positions (all other positions:  $|t(65)| < 2.0$ ,  $p_{FDR} > 0.1$ ). When quantifying this pattern using a logistic regression model, we found significant effects of both condition ( $b_{cond} = 0.28$ ,  $SE = 0.11$ ,  $CI = [0.06, 0.49]$ ,  $p = 0.012$ ) and sequence position ( $b_{seq\_pos} = -0.20$ ,  $SE = 0.02$ ,  $CI = [-0.25, -0.16]$ ,  $p < 0.001$ ), as well as a marginal interaction ( $b_{int} = 0.07$ ,  $SE = 0.04$ ,  $CI = [0.003, 0.14]$ ,  $p = 0.059$ ). Although this interaction does not quite reach significance, it suggests that predictable action sequences afforded a marginally greater enhancement in order memory for items seen at the end versus the beginning of an event.

### ***Item memory is not impacted by engagement in predictable action sequences***

Finally, we turned to the critical question of how engaging in a familiar action sequence during encoding impacts detailed memory for the novel visual items from each event. During this item recognition test, participants had to discriminate between items that had been presented during one of their errands (OLD items) and similar lures (NEW items); Figure 1.6. We computed d-prime as a measure of recognition accuracy, adjusted for extreme values following (Hautus, 1995). Using this metric, we found no difference in item memory as a function of whether an item (or its matched pair) had been encountered in the predictable versus random store ( $t(51) = -0.29$ ,  $p = 0.78$ ,  $CI = [-0.33, 0.26]$ ,  $d = 0.05$ ,  $CI_d = [-0.31, 0.41]$ ); Figure 1.6a. We also computed a corresponding Bayes Factor ( $BF$ ) for this analysis, which allows us to more directly assess evidence for the null hypothesis (i.e., that predictable action sequences had no effect on item recognition). To this end, we calculated a close approximation of  $BF_{null}$  based on the Bayesian Information Criterion (Brydges & Bielak, 2020; Rouder et al., 2009). We obtained a  $BF_{null} = 6.44$ , which indicates moderate evidence in favor of the hypothesis that there is no difference in

item recognition performance as a function of encoding condition.



**Figure 1.6. Item recognition memory test results for Experiment 3.** The topmost bar plot (**a**) shows the difference in  $d'$  (a common measure of recognition accuracy) for items from the predictable versus random store. Panel (**b**) shows the rate of each response type during the recognition test, split by condition. (FA = false alarm; CR = correct rejection). Examples of OLD images and corresponding NEW lures are shown in (**c**). Error bars indicate within-subject standard errors, and dots represent individual participants.

To examine item recognition at a more granular level, we next split trials into four different response bins: hits (responding OLD to an OLD item), misses (responding NEW to an OLD item), false alarms (responding OLD to a NEW item), and correct rejections (responding NEW to a NEW item). Participants were just as likely to make each type of response for items associated with the predictable store versus those associated with the random store — confirming that the predictable structure of one's motor behavior during encoding had no significant impact on the visual encoding of individual items

(hits:  $t(52) = -0.69$ ,  $p = 0.50$ ,  $CI = [-0.06, 0.03]$ ,  $d = 0.10$ ,  $CI_d = [-0.18, 0.38]$ ; false alarms:  $t(52) = -0.51$ ,  $p = 0.61$ ,  $CI = [-0.09, 0.06]$ ,  $d = 0.10$ ,  $CI_d = [-0.30, 0.51]$ ); Figure 1.6b. We also confirmed that no differences emerged when looking only at high- or low-confidence responses, or when examining memory as a function of sequence position (Appendix A, Figure 3). Ultimately, these null effects suggest that the impact of familiar action sequences on visual order memory is unlikely to stem from increased attention to and/or encoding of those novel items, but rather through a mechanism that selectively supports linking information together across time.

### 1.4.3 Discussion

In Experiment 3, we again showed that engaging in a familiar action sequence enhances memory for the order of novel, temporally-coincident events. Conversely, familiar actions did not affect memory for the details of individual items. Although null effects must be interpreted cautiously, these findings suggest that participants did not simply allocate more attention to novel stimuli in the presence of prior knowledge. Instead, we argue that their traversal through a known behavioral routine allowed novel items to be embedded within a stable, pre-existing memory representation, which specifically supports memory for temporal relationships.

## 1.5 General Discussion

Across three experiments, we find that novel visual item sequences were better remembered when they were encountered during the execution of a familiar motor sequence. This work significantly extends previous studies showing enhanced learning of new information from the same modality and/or conceptual space as existing knowledge (e.g., Anderson et al., 1978; Bein et al., 2019; King et al., 2019; Reder et al., 2016; van Kesteren et al., 2013). Specifically, here we show that engaging in highly familiar action sequences — which are ubiquitous in everyday life and yet underappreciated in existing literature — is sufficient to boost one-shot learning of unrelated, visual items seen simultaneously.

This work also deepens our understanding of how motor and episodic memory systems interact. Existing research has shown that motor sequence learning can benefit performance on an asynchronous

and unrelated memory task (Mosha & Robertson, 2016; Mutanen et al., 2020). In this previous work, the learning of a finger-tapping sequence enhanced participants' ability to recall a repeatedly-studied word list, and vice versa. Here we demonstrate that engagement in familiar behaviors can also enhance memory for episodic sequences encoded *at the same time*. This distinction is significant, given that our everyday actions are frequently coincident with exposure to other stimuli. Interestingly, in this prior work, cross-task learning transfer only occurred when finger-tapping sequences and word lists shared a high-level structure, such that the order of fingers in the motor task mirrored the order of semantic categories in the word list. The authors subsequently argued that this learning transfer occurred via "memory leaks," whereby abstract, sequential information learned in one context can be shared across multiple memories (Robertson, 2022). In our task, item and action sequences did not share this kind of abstract, high-level structure — which, according to their framework, is a necessary prerequisite for memory leaks. Our findings instead demonstrate that an abstract correspondence need not be present for cooperation between memories to occur, and that the simultaneous co-activation of motor and episodic sequence representations improves temporal memory — whereas in paradigms involving separate tasks, a more abstract shared structure may be necessary for integration across asynchronous sequence memories.

We further suggest that familiar action sequences benefit order memory by activating a holistic representation of the learned event structure during encoding (Figure 1.1b), within which novel item representations can become embedded (Gilboa & Marlatte, 2017; McClelland, 2013). Because this "temporal scaffold" specifically represents how events unfold across time, we predicted and found that order memory, and not item memory, would benefit from this structure. This putative explanation is rooted in work on the neurocognitive bases of sequence memory, which finds that memory for temporal links between items is enhanced when they belong to similar mental contexts (Clewett, Gasser, et al., 2020; DuBrow & Davachi, 2013; Heusser et al., 2018). Because repeated exposure to sequential information can increase the stability of sequence representations in the brain (Hsieh et al., 2014; Kalm et al., 2013), we hypothesize that well-learned action sequences bolster the stability of participants' mental context while encoding novel items, which then enhances temporal binding.

Similarly, the stable representation afforded by the familiar action sequence may also benefit order memory retrieval. Successful episodic recall is thought to be accompanied by reactivation of a memory's broader context, which in turn promotes access to other memories nearby in time (Howard & Kahana, 2002; Manning et al., 2011; Polyn et al., 2009). In our task, during retrieval of predictable events, participants may reactivate their memory for the known action sequence to guide recovery of the corresponding item sequence. In future work, we aim to clarify the precise neural mechanisms by which action sequences support temporal memory.

We also considered an alternative hypothesis, the attentional resource hypothesis — or the idea that prior knowledge about one's behavior diverts more attention to new information (Diana & Reder, 2006; Reder et al., 2016) — as the potential mechanism driving our temporal memory effects. However, if enhanced memory in our experiments stemmed from a broad increase in goal-directed attention during predictable events, we would expect benefits for item in addition to order memory. Given that item memory was unaffected by familiar actions, differences in attention are unlikely to explain our results.

It is perhaps surprising that familiar action sequences did not robustly facilitate spatial context memory. If participants remember the order of novel items and of predictable aisle responses, they could theoretically infer item-aisle pairings. However, the lack of clear spatial memory effects in our experiments suggest that participants did not adopt that strategy here. Further, although spatial context memory was enhanced for items in the *first* position of predictable versus random effects across Experiments 1 and 2, this benefit alone could not explain our finding of enhanced temporal memory for the *complete* order of items within predictable events. It could be that the format of the spatial memory test — where participants made spatial context judgments for one item at a time — hindered access to other items in the event and their temporal position. It is also possible that in our paradigm, spatial information was simply not salient enough to be integrated within item sequences.

It is important to acknowledge that in our task, we cannot be certain that the execution of motor behaviors during encoding is *necessary* for temporal sequence memory enhancements to occur. Instead, it is also possible that predictability within an event in general — e.g., expectations about where to-be-

encoded items will appear in space — is itself enough to enhance temporal memory, even when no responses are made. In the current experiments, we chose to have participants execute the learned sequence to ensure that this predictable event structure was made sufficiently salient during simultaneous encoding. Without this demand, there is the concern that making the action sequence irrelevant to ongoing episodic encoding may reduce its ability to impact memory processing, given prior work showing that contextual features of an event must be task-relevant in order to influence memory (e.g., Ecker et al., 2007). However, future work examining the effects of predictable sequential structure *without* any motor execution will be necessary to establish to what degree sequential action execution behavior is an essential prerequisite to our observed effects.

Ultimately, these experiments demonstrate how simple behavioral routines can support the formation of temporally-intact episodic memories. A substantial portion of our experiences are action-centered. Investigating how these actions impact concurrent memory processes, then, is critical to our comprehension of how learning occurs in real-world, naturalistic environments. Additional work will be necessary to establish the limitations of our effects. It could be, for example, that when behavioral routines become *too* automatic, they instead lead to disengagement from external stimuli (e.g., “zoning out” during a familiar drive). Nevertheless, the demonstration that cross-modal facilitation of episodic memory can occur is an important and novel contribution to our understanding of how prior knowledge facilitates new learning, as well as how different memory systems cooperate to support learning behavior.

**Chapter 2: Predictable action sequences scaffold memory for temporal  
structure and modulate hippocampal and cortical encoding activity**

Camille Gasser & Lila Davachi

**\*\* PLEASE NOTE:** The study presented in this chapter uses a very similar paradigm as the experiments discussed in Chapter 1. However, during the process of preparing this chapter/manuscript, we chose to rename the *random* condition to *variable*.

## B.1 Introduction

Prior knowledge about the world is essential for navigating everyday life. Such knowledge, gradually built and refined through past experience, shapes our expectations and can help us behave effectively in a variety of situations. A major focus of research on prior knowledge has been its effects on memory for new experiences (Alba & Hasher, 1983; Brod et al., 2013; Gilboa & Marlatt, 2017; Graesser & Nakamura, 1982; van Kesteren et al., 2012). Much work has shown that novel information is better remembered when it is consistent with our existing knowledge (Anderson et al., 1978; Bransford et al., 1972; Brewer & Treyens, 1981; DeWitt et al., 2012; Reder et al., 2013). This memory benefit has been explored with diverse kinds of knowledge structures, including general semantic knowledge (Bein et al., 2014; van Kesteren, Beul, et al., 2013; van Kesteren et al., 2020; van Kesteren, Rijpkema, et al., 2013), learned spatial environments (Guo & Yang, 2023; Sommer, 2017; Tse et al., 2007; van Buuren et al., 2014), familiarity with people or characters (Bein et al., 2019, 2020; Bellana et al., 2021; Kole & Healy, 2007; Z. X. Liu et al., 2017; Raykov et al., 2020, 2021), academic subject matter (van Kesteren et al., 2014, 2018), and event schemas or scripts (Bonasia et al., 2018; Masís-Obando et al., 2022).

Critically, this body of research suggests that prior knowledge might make new encoding more robust by enabling the formation of a memory representation that integrates new details with existing knowledge. However, several open questions about this process remain. For one, the majority of previous studies have focused selectively on the learning of novel information that is meaningfully related to or from the same domain as prior knowledge, leaving it unclear whether prior knowledge can also benefit memory for unrelated stimuli. Past work has also rarely examined knowledge structures that are explicitly temporal in nature (i.e., that describe how situations unfold across time), or investigated how the specificity of prior knowledge — that is, whether it affords predictions that are precise or general — impacts its ability to scaffold memory.

We previously conducted a behavioral study that explored the effects of specific motor sequence knowledge on simultaneous episodic sequence encoding (Gasser & Davachi, 2023; Chapter 1). We used a variant of the Ezzyat-DuBrow-Davachi (EDD) paradigm for this purpose, in which participants encode sequences of stimuli that are organized by contextual features into discrete episodic “events” (Buonomano et al., 2023; DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011; Heusser et al., 2018). Critically, our task additionally required participants to execute sequences of simple motor actions as part of each encoding event. Across three separate experiments, we found that the execution of a predictable and well-learned action sequence during encoding (relative to a random sequence of actions) selectively enhanced memory for the temporal order of objects within an event. This finding suggests that specific motor sequence knowledge can scaffold episodic memory for an unrelated sequence of items. Now, in the present study, we seek to uncover the neural mechanisms driving this memory facilitation.

Previous theories have argued that prior knowledge-mediated learning depends on bidirectional communication between the hippocampus and neocortex — most notably, the medial prefrontal cortex (mPFC) (Gilboa & Marlatte, 2017; McClelland, 2013; Preston & Eichenbaum, 2013; van Kesteren et al., 2012). Such cross-regional coupling is thought to allow novel hippocampal engrams to become rapidly integrated with existing cortical knowledge stores. Evidence for this framework comes from empirical findings in both animals and humans. Influential studies in rodents suggest that novel learning within a familiar spatial schema accelerates the process of systems consolidation, whereby new memory traces become independent of the hippocampus and assimilated within cortical regions such as mPFC (Takeuchi et al., 2022; Tse et al., 2007, 2011). In humans, neuroimaging studies indicate that the encoding of prior knowledge-related information is associated with increased univariate activity in mPFC (Bonasia et al., 2018; Brod & Shing, 2018; Z. X. Liu et al., 2017; Sommer et al., 2022; van Kesteren, Beul, et al., 2013; van Kesteren et al., 2014), as well as altered functional connectivity between mPFC and the hippocampus (Bein et al., 2014; Bonasia et al., 2018; Sommer, 2017; van Kesteren et al., 2014; van Kesteren, Fernandez, et al., 2010). Interestingly, other cortical hubs of the default mode network — such as posterior medial cortex (PMC) — show similar evidence for modulation by prior knowledge, such as

increased univariate activity during knowledge-congruent encoding (Bonasia et al., 2018; Maguire et al., 1999; Raykov et al., 2020; Sommer, 2017; Sommer et al., 2022).

The findings above dovetail with additional research indicating that mPFC and PMC represent structured knowledge during online experience (Baldassano et al., 2018; Fairhall & Caramazza, 2013; Raykov et al., 2021; Stawarczyk et al., 2021). In recent years, the strength of cortical knowledge representations during learning has also been connected to the success of subsequent memory. For example, Masís-Obando et al. (2022) found that successful encoding of narrative stimuli that followed common event schemas (e.g., going to a restaurant) was associated with the activation of abstract schema representations in anterior mPFC and narrative-specific representations in both mPFC and PMC. Relatedly, Guo and Yang (2023) showed that the representation of learned spatial structure in the lateral occipital cortex (LOC) was linked to improved encoding of novel associations within that structure. These two findings support the idea that cortical knowledge representations can scaffold the learning of new episodic content. In both studies, however, researchers did not completely disentangle the neural representation of the knowledge structure itself from those of individual stimuli: specifically, Masís-Obando et al. (2022) extracted schema representations by averaging brain activity patterns across multiple to-be-encoded narratives with shared schematic structure, whereas Guo and Yang (2023) did so by measuring activity while participants were cued with individual objects to retrieve a learned spatial schema. These tasks also did not assess how prior knowledge representations can scaffold memory for sequential information in particular. We address both of these points in the current work.

If prior knowledge-mediated learning is supported by the active representation of that knowledge during encoding, this mechanism might interact with other kinds of representations that are maintained during online experience: namely, the representation of context. There is an emerging literature showing that the temporal structure and organization of memories is modulated by stability and change in one's context (e.g., their physical environment or mental state). Specifically, information encoded in the same context is likely to be integrated together into a unified event memory, whereas items from distinct contexts are typically separated in memory (Clewett et al., 2019; Clewett & Davachi, 2017; DuBrow &

Davachi, 2013; Shin & DuBrow, 2021; Zacks, 2020). Previous studies have found evidence that the stability of neural activity patterns across time — a putative index of context stability — is linked to temporal memory performance. In particular, greater similarity between item representations in the hippocampus and LOC during encoding has been associated with enhanced memory for temporal order (DuBrow & Davachi, 2014), as well as compressed judgments of temporal distance (Ezzyat & Davachi, 2014). Building on this literature, one possibility is that prior knowledge may modulate this temporal memory mechanism by increasing context stability. In other words, we hypothesize that motor sequence knowledge will enhance temporal order memory by increasing the stability of context representations as encoding proceeds, ultimately facilitating temporal integration between items.

To investigate this idea, we used fMRI during the performance of our previously-validated behavioral paradigm (Gasser & Davachi, 2023; Chapter 1). In this task, participants ran a series of “errands” in two different stores, each of which required them to visit a sequence of aisles (by making simple motor responses) while simultaneously encoding a novel sequence of object images. The two stores functioned as our two encoding contexts, which varied in the type of prior knowledge they invoked. While both stores were associated with a schema-congruent category of object stimuli (e.g., food stimuli in the grocery store), the “predictable store” additionally afforded specific motor sequence knowledge, such that participants visited the exact same sequence of aisles (via the same sequence of motor responses) during every errand. We first replicated our previous behavioral findings by showing that the execution of a predictable action sequence (relative to an unpredictable sequence) selectively bolstered temporal order memory for novel and unrelated object sequences. Then, we explored how this learned motor structure (i.e., the “predictable action scaffold”) shaped brain activity during encoding. As part of this exploration, we began by probing the neural representation of our two encoding contexts, asking where and how the predictable action scaffold was represented in the brain. We then evaluated the stability of neural activity throughout events as a putative mechanism of successful temporal sequence encoding. Finally, we tested whether the strength with which the predictable action scaffold was represented during encoding bolstered the stability of neural activity across time.

For all analyses, we targeted a diverse set of brain regions that we hypothesized would represent the action scaffold and/or temporal context in our task. These regions included: (1) mPFC and (2) PMC, given their established role in the representation of prior knowledge and its effects on memory (Bonasia et al., 2018; Gilboa & Marlatte, 2017; Masís-Obando et al., 2022; Sommer et al., 2022); (3) premotor cortex, which is known to represent motor sequence information (Crowe et al., 2014; Kornysheva & Diedrichsen, 2014; Wiestler & Diedrichsen, 2013; Yokoi & Diedrichsen, 2019); (4) the hippocampus, an integral locus of episodic memory function that is sensitive to temporal and sequential structure (DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014; Hsieh et al., 2014); and (5) the lateral occipital cortex (LOC), which represents visual objects (Grill-Spector et al., 2001; Malach et al., 1995) and has also been found in previous work to track context stability (Ezzyat & Davachi, 2014, 2021).

## B.2 Methods

### B.2.1 Participants

For the current study,  $N = 33$  healthy young adults were recruited from the Columbia University community. All participants were fluent English speakers, reported having normal or corrected-to-normal vision, and indicated that they were not taking psychoactive medication at the time of the study. From this sample, three participants were excluded prior to analysis due to failure to learn the predictable motor sequence before the encoding task ( $N = 2$ ) or excessive motion during fMRI scans (mean frame-wise displacement  $> 0.2$  mm across the majority of scans;  $N = 1$ ). Following these exclusions, we were left with a final sample of  $N = 30$  (21 female; 18–35 years old;  $M_{age} = 23.9$ ;  $SD_{age} = 4.2$ ). All participants provided informed consent and were compensated at a rate of \$30/hour for time spent in the MRI scanner and \$12/hour for behavioral tasks outside of the scanner. All study procedures were approved by the Institutional Review Board at Columbia University.

### B.2.2 Stimuli

Stimuli consisted of colorful, high-resolution images of objects of animals (e.g., cat, dolphin) and foods (e.g., donut, milk). For each unique object, we collected two distinct exemplar images, which we

used to construct two matching sets of stimuli (set A and set B). These sets contained the same collection of object types, but different images. There were 288 unique objects (144 animals; 144 foods) within each set. Each participant saw objects from one of the two sets during encoding, while the other was used to provide lures for the object recognition test. For the training task, we also collected 24 abstract fractal images. All images were uniformly resized and had their backgrounds removed via Adobe Photoshop CC (2021 release).

### B.2.3 Task design

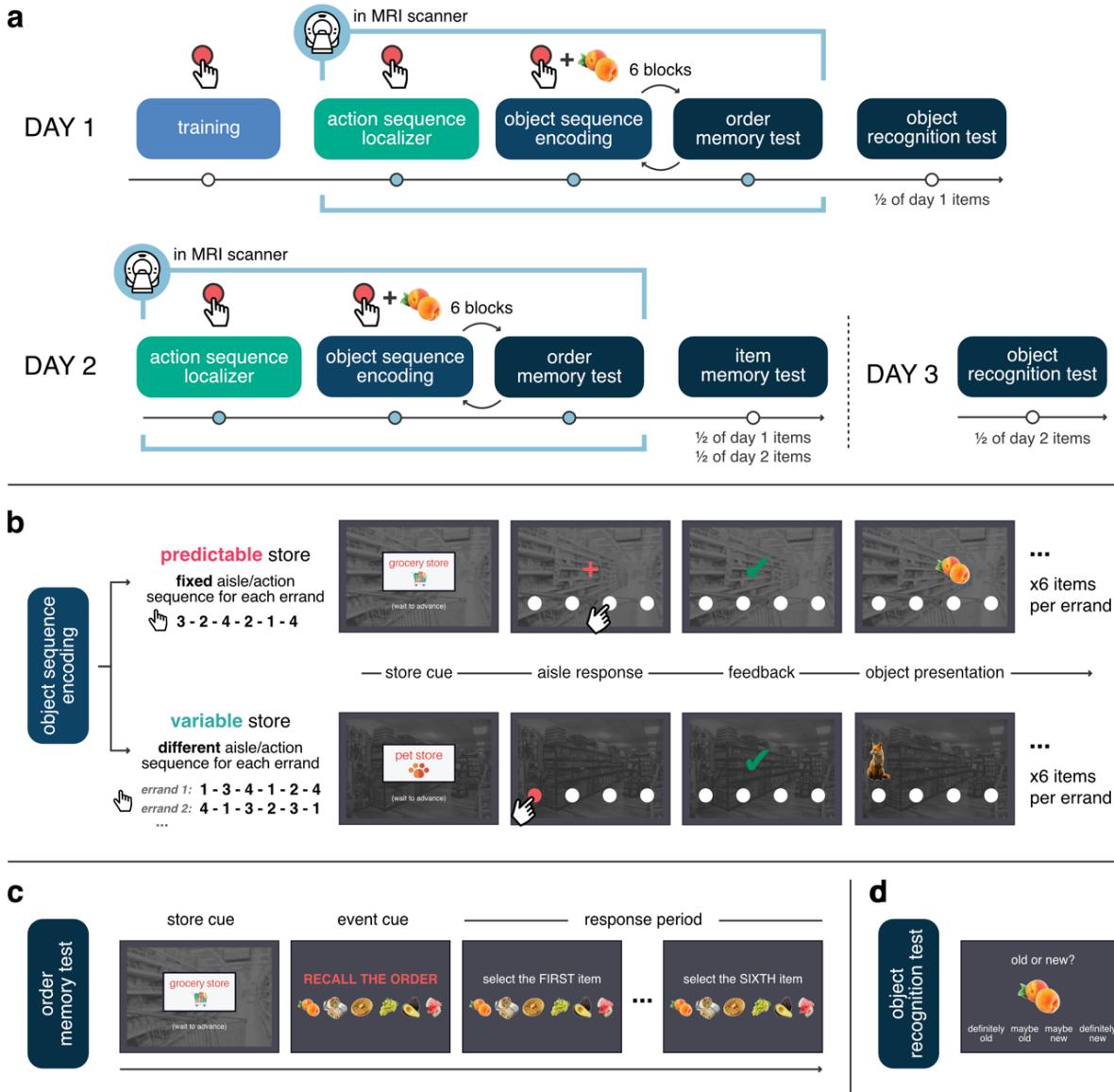
**Overview.** The design of the current experiment was adapted from our previously-published behavioral study (Gasser & Davachi, 2023; Chapter 1). Participants' task was to run errands through two different stores: the pet store and the grocery store. As part of each errand, participants visited a sequence of six aisles in the store by executing simple motor responses (i.e., button presses) while simultaneously "collecting" (i.e., encoding) a novel sequence of visual objects. The two stores constitute our two conditions, or encoding contexts: in the *predictable* store, the sequence of aisles that participants visited (i.e., the motor action sequence they executed) always followed the same, fixed pattern. In the *variable* store, this aisle sequence was random and changed with each errand. (Note that in Chapter 1, the equivalent encoding condition was referred to as *random* instead of *variable*.) Critically, the objects encoded within both stores were always novel. Following each set of errands, participants were tested on their memory for the objects collected within each store.

The entire study took place across three consecutive days, each approximately 24h apart. On Day 1, participants first learned the predictable aisle sequence during a behavioral session outside of the MRI scanner ("training"). They then completed several blocks of the errand-running task (or "object sequence encoding") inside of the scanner on both Day 1 and Day 2. Memory for novel object sequences encoded during these errands was assessed during multiple memory tests administered both within and outside of the MRI scanner. All tasks are described in detail throughout the following sections and summarized in Figure 2.1.

**Training.** At the start of the experiment, prior to beginning their errands, participants were

trained on the motor sequence they would perform each time they entered the predictable store. The training was organized into three study-test cycles, following the same structure used in Chapter 1 (Gasser & Davachi, 2023). The action sequence consisted of six button presses which corresponded to four possible aisles, such that two out of the four aisles were visited twice as part of the same sequence (see *Aisle sequence generation*). During study blocks, participants practiced executing the predictable motor sequence. Each study repetition began with a banner cue displaying the identity of the upcoming store (for 3s), followed by a screen displaying the four aisles (four side-by-side circles) with a static, greyscale image of the store interior in the background. Participants were initially guided to execute the six-item motor sequence via simple visual cues, in which the circle associated with the upcoming aisle turned red to prompt a response. Responses were made via button presses on a laptop keyboard, specifically using the index and middle fingers on each hand. Each response had a 6s time limit and was followed by a 0.5s delay before the next aisle cue appeared. Only correct responses were accepted. Each study block was followed by a brief distractor task lasting approximately 20s (see *Distractor task*) before the subsequent test block began. During this test, participants had to recreate the predictable store's aisle/action sequence from memory. Feedback was provided for 0.5s after each motor response (green checkmark for correct responses; red X for incorrect responses), as well as at the end of the entire test block (out of 100%).

As mentioned above, the training task included three study and three test blocks. In the first two training blocks, participants were cued to execute the predictable motor sequence five times in a row. However, during the third study block, the predictable sequence was interleaved with the execution of two cued *variable* motor sequences that were also cued. In addition, after each motor response during this block, an abstract fractal image appeared in the visited aisle (i.e., above the aisle circle) for 2.5s. Importantly, the structure of this final study block foreshadowed the structure of the upcoming encoding/errand task, in which participants would visit both the predictable and variable stores in an intermixed fashion and view novel objects after each aisle response. During all test blocks of the training task, participants were asked to recreate the predictable aisle/action sequence from memory during two consecutive test trials.



**Figure 2.1. Experiment design.** (a) Schematic of all tasks completed on all three days of the experiment. (b) Illustration of the encoding event procedure in the predictable and variable conditions. In the predictable store, participants visited the same aisle sequence (i.e., executed the same sequence of motor actions) during every errand, which they had to retrieve from memory. In the variable store, participants executed a different aisle sequence on each errand, with responses explicitly cued. (c) Illustration of the order memory test that followed each block of errands. (d) Example trial from the object recognition task, where participants discriminated between previously-encoded objects and visually similar lures.

Of note, two participants in our sample indicated to the experimenter that they had misunderstood instructions during the final training study-test cycle by believing their goal was now to memorize aisle

sequences within the *variable* store. To resolve this confusion and ensure that these participants exited the training phase with intact memory for the predictable aisle sequence, we had them complete one extra study block (with the same structure as the first two study blocks) followed by one additional test block.

***Object sequence encoding (“errand-running”).*** The object sequence encoding task required participants to run errands, during which they “visited” a sequence of aisles (i.e., executed a series of motor actions) while simultaneously encoding a sequence of novel objects. Each encoding block included four errands/events: two in the predictable store and two in the variable store (order counterbalanced). Critically, during all predictable events, participants visited the same sequence of aisles as had been learned during training. During variable events, they followed a different aisle/action sequence for each errand (Figure 2.1b).

Each errand began with a banner depicting the identity of the upcoming store for 3s. Next, participants visited six aisles and collected six objects in a similar manner as during the final training study block. As in Experiments 1 and 3 from Chapter 1 (Gasser & Davachi, 2023), the process of visiting aisles varied slightly for predictable versus variable events: in the predictable store, participants had to *use their memory* to visit the correct sequence of aisles, whereas in the variable store, they saw explicit *cues* (i.e., the aisle circle turning red) indicating which response should be made next. All responses were self-paced, up to a maximum time limit of 2.5s. Although this procedure results in slightly different task demands across the two conditions, our previous behavioral work established that this design feature does *not* explain resulting differences in memory for objects from predictable versus variable events, as differences in memory were evident even when explicit cues were provided in both conditions (Gasser & Davachi, 2023; Chapter 1). We chose this task structure for the current study because requiring participants to execute the predictable motor sequence from memory not only produced numerically stronger memory effects in our previous work, but also better resembles the structure of everyday life, where people must often retrieve existing knowledge to guide their behavior.

In both store conditions, each aisle/action response was immediately followed by feedback (green checkmark for correct responses; red X for incorrect responses), which remained on the screen for 0.5s. A

novel object then appeared in the visited aisle for 2.5s, followed by a jittered fixation period of 0.5-2s before the onset of the next aisle response period. Throughout the duration of each errand, a static greyscale image of the store's interior remained in the background as a constant reminder of the current encoding context (Figure 2.1b). Objects collected during errands were drawn from semantic categories consistent with each store's identity, such that animals were seen in the pet store and food objects were seen in the grocery store. Aisle responses were made with an MRI-compatible button box, and participants were instructed to use the same finger-response mapping as they had during the training task. Each errand was followed by a jittered interval of 12-15s of an active baseline task (see *Distractor task*). There were 12 object sequence encoding blocks in total throughout the experiment, with six blocks on Day 1 and six on Day 2 (Figure 2.1a). All encoding blocks took place in the MRI scanner.

**Order memory test.** After each object sequence encoding block, participants completed an order memory test for the most recent set of errands (Figure 2.1c). Each test trial began with a cue banner indicating the identity (predictable or variable) of the to-be-tested store for 3s, before the presentation of a visual array containing all six objects from a single errand arranged horizontally across the screen (with their spatial positions randomized). This object array remained onscreen for 4.5s beneath a prompt stating “recall the order.” During this part of the trial, participants were instructed to try to remember the order of the indicated objects, but to withhold any responses. They then were prompted to select the objects in order one at a time. Responses were self-paced, and participants had up to 30s to select all six of the objects per event. After each test trial, participants completed an active baseline task (see *Distractor task*) for 12-15s. There were 12 order memory blocks in total (six on Day 1, six on Day 2), each of which tested memory for four errands (in a randomized order).

**Object recognition test.** Although not the main focus of this manuscript, we also assessed recognition memory for individual objects encoded during errands (Figure 2.1d). This recognition test took place outside of the MRI scanner on all three experiment days. Objects in the test were divided into two delay conditions, such that memory for half of all objects was probed on the same day as encoding (“immediate memory”), while the other half was probed on the day after (“delay memory”); see Figure

2.1a. This division resulted in 72 test trials on Days 1 and 3 and 144 trials on Day 2. During each trial, participants saw an object and selected whether it was “definitely old” (i.e., had appeared in one of their errands), “maybe old”, “maybe new”, or “definitely new.” All new images (i.e., lures) belonged to the same category as an object studied during encoding, such that participants had to remember the specific perceptual details of each object image in order to perform well. The order of trials was pseudo-randomized to ensure that different trial types were spread out evenly throughout the task. Specifically, we ensured that each third of the test period contained an approximately even number of: objects from predictable versus variable stores; objects from early versus late blocks of the encoding task (within each day); objects tested immediately versus after a delay; and old versus new objects. Trials were self-paced with a time limit of 8s. Importantly, note that all (old) objects that appeared during the recognition test had already been tested as part of the order memory test. As such, object recognition performance was likely influenced by the additional encoding opportunity that order memory trials provided.

**Action sequence localizer.** Prior to the start of encoding on Days 1 and 2, participants completed one block of the action sequence localizer task (Figure 2.1a). This task followed a highly similar structure as errands during encoding, with the primary difference being that *no novel objects* were presented following each aisle/action response. As such, this task allowed us to obtain an independent measure of brain activity as participants executed predictable (and variable) motor sequences. Trials in the localizer task began with a 3-s cue banner depicting the upcoming store’s identity before cuing the participant to execute a 6-object motor sequence. Participants were given 2.5s to make each response, and a fixed interval of 3s separated the onsets of successive aisle cues. Each run of the localizer included four repetitions of the predictable sequence, as well as four distinct variable sequences — with the order of predictable versus variable sequences randomized. All aisle responses in both conditions were explicitly cued and only correct responses were accepted, in order to guarantee that participants executed the correct sequence of actions during each trial. As with encoding events and order memory trials, distractor task periods lasting 12-15s followed each localizer trial.

**Distractor task.** Participants performed a simple distractor task between errands during object

sequence encoding, between trials during the order memory test, and between trials (i.e., sequence repetitions) in the action sequence localizer task. During this task, a stream of single-digit numbers was presented one at a time on the screen, and participants were instructed to press a button whenever an even number appeared. Each number was presented for 1.25s, followed by a 0.5s inter-trial-interval.

**Aisle sequence generation.** Each aisle sequence was composed of six consecutive actions corresponding to four possible response options (i.e., aisles). Thus, two of the four aisles in each store were visited twice as part of each sequence. Participants were randomly assigned one of two possible predictable aisle sequences, which they followed every time they visited the predictable store, (i.e., during training, the action sequence localizer, and object sequence encoding). These predictable sequences were drawn directly from our previous study, which describes the process of sequence generation in greater detail (Gasser & Davachi, 2023; Chapter 1). Each participant also saw a total of 34 unique variable sequences throughout the experiment (2 during training, 8 during the action sequence localizer, and 24 during encoding). Variable sequences had the same structure and constraints as predictable sequences (e.g., 6 consecutive responses, 2 aisles revisited, no aisles repeated back-to-back). Critically, they were constructed in order to minimize their similarity to the predictable sequence, as well as to minimize the overlap between different variable sequences. The specific criteria used to create these variable sequences is also detailed in Gasser & Davachi (2023), with one minor difference: namely, due to the larger number of unique variable sequences used in the present study (as compared to our previous behavioral experiments), we removed the requirement that the first aisle in each variable sequence could not be revisited. Importantly, our previous work showed that aisle revisiting did not interact with the effects of predictable versus variable motor sequences on memory.

#### B.2.4 MRI acquisition

Scanning was performed at the Zuckerman Institute at Columbia University using a 3T Siemens Magnetom Prisma scanner with a 64-channel head coil. At the start of each scanning day (Days 1 and 2), structural images were collected using a T1-weighted magnetization-prepared rapid-acquisition gradient echo (MPRAGE) sequence (1mm isotropic voxels), as well as a T2-weighted echo-planar sequence

(0.7mm isotropic voxels). Whole-brain functional images were acquired using a multiband echo-planar imaging (EPI) sequence (repetition time = 1.5s, echo time = 30ms, in-plane acceleration factor = 2, multiband acceleration factor = 3, voxel size = 2mm isotropic), with 69 oblique axial slices acquired in an interleaved order. All slices were tilted 14° relative to the AC-PC line. There were seven functional scans during each scanning day: one block of the action sequence localizer task and six blocks of the errand task (note each set of encoding and order memory test blocks occurred during the same scan). At the end of each scanning session, we also collected gradient echo (GRE) field maps using the same slice prescription as the EPI scans, in order to improve co-registration between anatomical and functional images.

### B.2.5 fMRI preprocessing

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 20.2.3 (Esteban et al., 2019, 2022; RRID:SCR\_016216), which is based on *Nipype* 1.6.1 (Gorgolewski et al., 2011; RRID:SCR\_002502).

**Anatomical data preprocessing.** A total of 2 T1-weighted (T1w) images were found within the input BIDS dataset. All of them were corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.3.3 (Avants et al., 2009; RRID:SCR\_004757). The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTS as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR\_002823, Zhang et al., 2001). A T1w-reference map was computed after registration of 2 T1w images (after INU-correction) using mri\_robust\_template (FreeSurfer 6.0.1, Reuter et al., 2010). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR\_001847, Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR\_002438, Klein et al., 2017). Volume-based spatial normalization to two standard spaces (MNI152NLin2009cAsym, MNI152NLin6Asym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both

T1w reference and the T1w template. The following templates were selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c [Fonov et al., 2009, RRID:SCR\_008796; TemplateFlow ID: MNI152NLin2009cAsym], FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model [Evans et al., 2012, RRID:SCR\_002823; TemplateFlow ID: MNI152NLin6Asym].

***Functional data preprocessing.*** For each of the 14 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIprep. A B0-nonuniformity map (or fieldmap) was estimated based on a phase-difference map calculated with a dual-echo GRE (gradient-recall echo) sequence, processed with a custom workflow of SDCFlows inspired by the epidewarp.fsl script and further improvements in HCP Pipelines (Glasser et al., 2013). The fieldmap was then co-registered to the target EPI (echo-planar imaging) reference run and converted to a displacements field map (amenable to registration tools such as ANTs) with FSL's fugue and other SDCFlows tools. Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (D. N. Greve & Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, Jenkinson et al., 2002). The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into several standard spaces, correspondingly generating the following spatially-normalized, preprocessed BOLD runs: MNI152NLin2009cAsym, MNI152NLin6Asym. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIprep. Several

confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al., 2014) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al., 2002). FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by Power et al., 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor, Behzadi et al., 2007). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's aseg segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardized DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e.

head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using mri\_vol2surf (FreeSurfer).

Many internal operations of *fMRIprep* use *Nilearn* 0.6.2 (Abraham et al., 2014, RRID: RRID:SCR\_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIprep*'s documentation.

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

## B.2.6 fMRI analyses

After preprocessing, all fMRI runs were smoothed with a 5mm FWHM (full-width at half maxima) Gaussian kernel. Note that although representational similarity analyses are sometimes performed on unsmoothed data, recent work has shown that moderate amounts of smoothing (i.e., 1-4 times the voxel size) can improve the sensitivity of multivariate pattern analyses (Hendriks et al., 2017). All fMRI analyses were performed using a combination of FSL (version 6.0.4; Jenkinson et al., 2012), python (ver. 3.9.10), and R (ver. 4.3.3; <http://www.r-project.org>), with all general linear models (GLMs) implemented by FSL's FEAT (Woolrich et al., 2001).

**Regions of interest.** For the majority of analyses in this manuscript, we adopted an *a priori* region of interest (ROI) approach. ROIs corresponding to medial prefrontal cortex (mPFC) and posterior medial cortex (PMC) were extracted from the Schaefer parcellation (with 400 parcels; Schaefer et al., 2018) by combining all parcels with the labels “PFCm” and “pCunPCC”, respectively. We also defined a region corresponding to the premotor cortex using the Human Motor Area Template (HMAT) atlas (Mayka et al., 2006) by combining dorsal and ventral subregions. Finally, participant-specific ROIs corresponding to the hippocampus and lateral occipital cortex (LOC) were created using FreeSurfer's segmentation tool

(Fischl, 2012). All ROIs were bilateral and were aligned to each participant's native T1 space prior to further analysis.

***Neural similarity across encoding events.*** When investigating how predictable motor structure is represented in the brain, we began by quantifying the similarity of activity patterns across encoding events. If participants have activated a neural representation of the predictable action scaffold, we should see greater consistency in the activity patterns evoked during all predictable events relative to variable events, where the learned motor structure is irrelevant. To test this idea, brain activity patterns for all individual encoding events were derived using a set of GLMs applied to each encoding run. Each GLM included a set of 8 task regressors corresponding to the store cue periods and object encoding periods for each of the 4 events in each block/run. Store cue periods were modeled as a 3-s boxcars. Each object encoding period was modeled as the sum of six consecutive boxcar functions, corresponding to the six objects/motor responses in the event. The duration of each boxcar in this summation extended from the onset of an aisle cue until the offset of the subsequent object's presentation (~3-5s). Given that the encoding and order memory test periods of each block occurred as part of the same scanning run, each GLM also contained regressors corresponding to phases of the order memory test (Figure 2.1c), including store cue periods, event cue periods (during which participants saw all 6 objects from an errand but did not make any responses), and response periods (during which they selected objects in order one at a time). Task regressors were again modeled as boxcars with the duration of the relevant task phase (i.e., 3s for store cues, 4.5s for event cues), with each response period regressor stretching from the beginning of the response period until participants selected the final/sixth object in the event (~10-20s). All task regressors were convolved with FSL's double-gamma hemodynamic response function (HRF). To account for participant head motion, each GLM also included six rigid-body motion parameters (i.e., translation and rotation along the X, Y, and Z axes), their temporal derivatives, and stick function regressors at timepoints detected as motion outliers (see *fMRI Preprocessing*). We also applied a high-pass filter with a 1/128 (0.008) Hz cutoff to reduce the impact of low-frequency fluctuations in the BOLD signal.

These GLMs provided us with a set of statistical maps representing activity during the object

encoding period of each event. For each event, a pattern of brain activity was extracted across voxels within each ROI and then z-scored. Finally, following a representational similarity approach (Kriegeskorte et al., 2008), we took each of these encoding patterns and computed its average neural similarity (i.e., Pearson correlation) to all other patterns from the same condition (*within-condition* similarity) and to all other patterns from the other condition (*across-condition* similarity). These similarity values were computed separately within each participant and day, and all correlation values were subjected to a Fisher R-to-Z transformation before further statistical testing. To account for the fact that temporal autocorrelation in the BOLD signal may inflate similarity between events belonging to the same run, only across-run pairs of events were included in this analysis.

If a particular brain region represents the predictable action scaffold, we would expect greater consistency between activity patterns evoked within predictable events relative to across contexts/conditions. Similarly, consistency between activity patterns for variable events might indicate that a region represents information specific to the variable store context. As such, we interpreted greater *within-condition* versus *across-condition* similarity within a given condition as evidence for the neural representation of that encoding context. In addition, we tested whether the strength or consistency of context representations differed between the predictable and variable stores by asking if the difference between *within-condition* and *across-condition* similarity was greater in one condition versus the other.

***Neural representation of the predictable action sequence.*** Next, we asked whether encoding activity in the predictable store contained information about the predictable action sequence. In order to extract a pattern of brain activity representing the predictable action sequence *without* containing any object information, we derived “sequence templates” from the action sequence localizer task (Figure 2.4). All localizer runs were modeled using GLMs, and each GLM contained 4 task regressors: predictable store cue periods, variable store cue periods, predictable motor response periods, and variable motor response periods. Store cue periods were modeled as 3-s boxcars. Motor response periods were modeled as boxcars that extended from the onset of the first aisle/action cue until the sixth and final response in the sequence (~15-16s duration). All regressors were convolved with FSL’s double-gamma HRF. Each

localizer GLM included the same set of confound regressors as used when modeling the encoding task.

From these localizer GLMs, we obtained statistical maps corresponding to brain activity during the execution of predictable and variable motor action sequences for each participant on each day of scanning. We then extracted patterns of activity from these maps within each of our ROIs. These patterns were z-scored across voxels, thereby removing the mean univariate signal. We refer to brain activity patterns associated with predictable versus variable motor sequences as “predictable sequence templates” and “variable sequence templates”, respectively. Note that for this analysis, we used sequence templates derived from the localizer task on Day 1, as this timepoint was most proximal to the successful learning of the predictable sequence. However, use of templates derived from Day 2 (when computing similarity to encoding events from that day) yielded similar results.

Next, in order to quantify the presence of sequence templates during encoding, we computed the neural similarity between each template and activity patterns for all encoding events (using the same encoding activity patterns as derived for the previous analysis); Figure 2.4. Our primary measure of interest was similarity between the *predictable sequence template* and *predictable encoding patterns* (P), as this measures the presence and strength of the learned action sequence representation during novel object encoding. We refer to this measure as “predictable template activation.” Critically, to establish the validity of this interpretation, we compared predictable template activation to three control measures of template similarity (Figure 2.4). First, we calculated similarity between the *predictable sequence template* and *variable encoding patterns* (C1). If our measure of interest (P) exceeds this first control, it establishes that the predictable sequence template is more strongly activated during predictable events — when the learned action sequence is relevant for behavior — than during variable events, when it is not. Next, we computed similarity between the *variable sequence template* and *variable encoding events* (C2). This measure addresses the fact that within both conditions, there is a high degree of similarity between participants’ experience during the sequence localizer and encoding tasks, in that both timepoints involve a consistent store identity/context and similar perceptual information (i.e., the same store interior image in the background). To the extent that our measure of interest (P) is greater than this control, we can assert

that the presence of a consistent motor sequence in the predictable condition enhances template activation above and beyond these factors. Finally, we calculated similarity between the *variable sequence template* and *predictable encoding patterns* (C3). If our measure of interest exceeds this control, we can argue that during predictable encoding, participants are not merely activating a general representation of the sequential task structure (i.e., the presence of six consecutive motor responses), in which case we would expect to see similar evidence for the variable sequence template. Instead, we can conclude that brain activity during encoding reflects information specific to the predictable motor sequence. These template similarity measures were computed within each ROI and then subjected to a Fisher R-to-Z transformation. We then used paired, two-sided t-tests to contrast predictable template similarity during predictable events (P) with all three controls.

***Within-event neural stability.*** A major goal of the present study was to understand how specific action sequence knowledge affects the formation of episodic sequence memories. Based on previous work, we focused on the stability of brain activity across time as a putative neural mechanism supporting temporal order memory. First, a new set of encoding GLMs was constructed to extract brain activity patterns corresponding to individual sequence positions within each encoding event. We adopted a Least Squares Separate (LSS) modeling approach (Mumford et al., 2012), whereby a separate GLM was constructed for each “sequence position” in each encoding event, resulting in 288 models (6 positions \* 4 events \* 12 blocks) per participant. Each model contained one regressor of interest corresponding to the presentation of a single object and its preceding motor/aisle response (and feedback period). Each of these regressors was modeled as a boxcar function (convolved with the HRF) lasting from the onset of the aisle response cue until the offset of the object stimulus (~3-5s). All other object presentations and task periods were modeled by 10 regressors of non-interest: predictable store cues, variable store cues, other predictable sequence positions, other variable sequence positions, and order memory test phases (store cues, event cues, and response periods) for predictable and variable events. These regressors also took the form of boxcar functions with durations corresponding to the length of the relevant task period. Each GLM included the same set of nuisance regressors included in our other task-based GLMs: six rigid-body

motion parameters, their temporal derivatives, and stick function regressors at timepoints detected as motion outliers (see *fMRI Preprocessing*).

From these GLMs, we obtained a set of statistical maps corresponding to brain activity associated with each individual sequence position (i.e., motor response + object stimulus) during the encoding task. Multivariate patterns were extracted from each map within each ROI and then z-scored. We then calculated “within-event stability” as the average correlation between brain activity patterns from all six sequence positions within an event. Greater stability values indicate that the pattern of neural activity throughout an event remained more consistent across time.

### B.2.7 Statistical analysis

The majority of our analyses were conducted using (two-sided) paired sample t-tests and multilevel linear regression models (see below). For paired contrasts involving reaction time data, we instead used Wilcoxon signed-rank tests, given that RT distributions often violate assumptions of normality made by parametric tests. For all tests, we report 95% confidence intervals (*CI*) and effect sizes: Cohen’s d for t-tests and Cliff’s delta for Wilcoxon signed-rank tests (Cliff, 1996). In the case of negative Cohen’s d or Cliff’s delta values, the sign was flipped such that reported effect sizes are always positive. For all fMRI analyses, FDR correction was used to account for multiple comparisons (Benjamini & Hochberg, 1995). Specifically, for each effect of interest, we corrected *p*-values across all of the ROIs considered in the analysis. All statistical tests were performed in R (ver. 4.3.3; <http://www.r-project.org>) using functions from the following packages: the built-in *stats* package (ver. 4.3.3), *tidyverse* (ver. 2.0.0; Wickham et al., 2019), *Rmisc* (ver. 1.5.1; Hope, 2022), *psycho* (ver. 0.6.1; Makowski, 2018), *effsize* (ver. 0.8.1; Torchiano, 2020), *lme4* (ver. 1.1-35.1; Bates et al., 2015), *lmerTest* (ver. 3.1-3; Kuznetsova et al., 2017), and *emmeans* (ver. 1.8.5; Lenth et al., 2018).

**Multilevel regression modeling.** For all regression models, we began by running a standard Frequentist model with random slopes for all main effects and random intercepts — all nested within each participant. Categorical predictors were effect-coded and continuous predictors were mean-centered. All Frequentist models were run using the *lme4* and *lmerTest* packages in R, with the *emmeans* package used

to evaluate the significance of simple slopes. When a model did not converge or resulted in a singular fit, we first re-ran a simplified version of the model by removing random slopes. We then ensured that results from this simplified model held when fitting a more complex Bayesian regression model with the full set of predictors. A Bayesian approach was used for this purpose because it often allows for better estimation of models with complex, hierarchical structures and/or with limited sample sizes. Moreover, when fitting these models with uninformative priors (as we do here), results from Bayesian models can be treated as conceptually analogous to their Frequentist counterparts. See Appendix B for details about the implementation of Bayesian regression models.

**Individual event exclusions.** Across all participants, two blocks of the errand task and subsequent order memory test (i.e., 8 events in total) were excluded *a priori* from all behavioral (and fMRI) analyses: one due to the participant's confusion over task instructions, and one because the scanning session was terminated early due to a fire alarm. We also excluded three individual events: two because technical issues caused them to be encoded twice (specifically: the projector malfunctioned, so the scan was restarted), and one because the participant failed to select any items during the order memory test trial.

## B.3 Results

### B.3.1 Training performance

At the start of the experiment, participants completed a training task to learn the aisle/action sequence they would follow whenever visiting the predictable store. To assess learning, we first examined reaction times (RTs) during the third study block of this task, during which participants were repeatedly cued to execute the predictable action sequence and several variable action sequences (see *Methods*). If participants successfully memorized the predictable action sequence, they should be faster to respond to aisle cues in the predictable condition, given that they could adaptively anticipate upcoming responses. Indeed, RTs were significantly faster during predictable versus variable action sequences (Wilcoxon signed-rank test:  $V = 45$ ,  $N = 30$ ,  $p < 0.001$ , Cliff's  $d = 0.451$ ). As part of this training task, participants also completed several test blocks, during which they attempted to execute the predictable action

sequence from memory. (Note that while the majority of participants performed three study-test blocks during training, two participants completed one additional block as a result of misunderstanding instructions during the third block — see *Methods*). Critically, by the final test trial, all participants reached 100% accuracy, indicating successful memorization of the predictable action sequence.

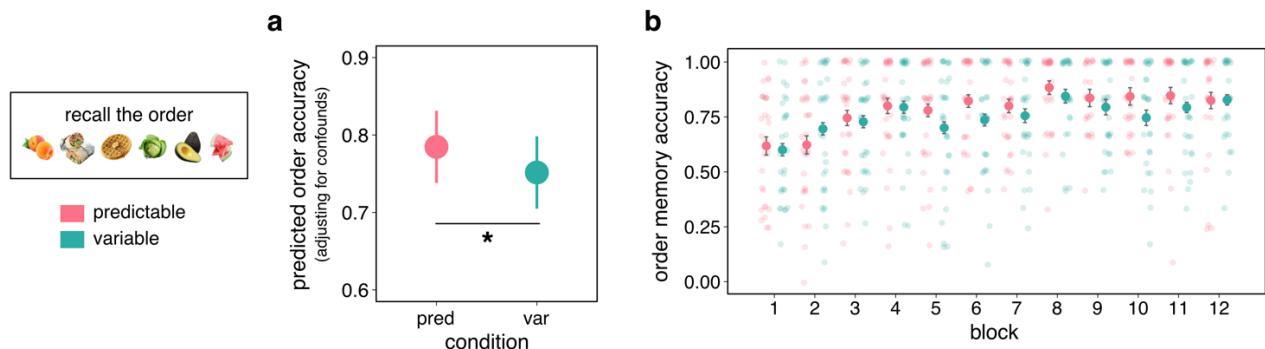
### B.3.2 Learning of action sequences persists throughout sequence localizer and encoding tasks

After training, participants executed the predictable action sequence throughout two different tasks: (1) the action sequence localizer, and (2) object sequence encoding (or “errand-running”); Figure 2.1a. During the action sequence localizer task, they were cued to visit sequences of aisles in the predictable and variable stores *without* viewing any items. Here, as during the final training study block, we found that participants were significantly faster to make predictable versus variable aisle responses (Wilcoxon signed-rank test:  $V = 2, N = 30, p < 0.001$ , Cliff’s  $d = 0.636$ ). This result suggests that participants could more effectively anticipate and deploy responses in the predictable (versus variable) store, indicating intact memory for the predictable action sequence.

During object sequence encoding, participants continued to execute aisle/action sequences while visiting the two stores — but now with the added task of encoding novel object sequences at the same time. Unlike the training and action sequence localizer tasks, where visual cues were provided to signal upcoming responses in both stores, the encoding task required participants to execute the predictable aisle/action sequence *using their memory*. In the variable store, explicit cues were still provided, given that this store had no consistent aisle sequence to remember. Despite this difference in task demands, the accuracy of aisle responses in both stores was near ceiling (predictable:  $M = 0.974, SD = 0.019$ ; variable:  $M = 0.980, SD = 0.015$ ) and did not differ significantly between conditions ( $t(29) = -1.164, CI = [-0.015, 0.004], p = .254, d = 0.306$ ). Next, we examined the speed of aisle responses during the encoding task. We found that participants were faster to execute actions in the variable store (where they were provided with explicit cues) than in the predictable store (where they had to use their memory to make responses), Wilcoxon signed-rank test:  $V = 359, N = 30, p = .008$ , Cliff’s  $d = 0.280$ .

### B.3.3 Execution of predictable action sequences scaffolds temporal order memory

Having established that participants learned and retained the predictable action sequence throughout our experiment, we next asked whether this learned motor structure facilitated temporal order memory for novel object sequences encoded in parallel. To test this question, we ran a linear regression model where the proportion of items recalled in the correct temporal position within an event was predicted by condition (predictable vs. variable), while controlling for several additional variables that we reasoned might interact with memory performance: block number, stimulus category, and the average of (log-transformed) aisle response RTs within the event. The model also included an interaction term between condition and block number, to account for the possibility that the effect of condition on order memory might evolve throughout the experiment. Note that to allow for model convergence, in this section we report results from regression models with only random intercepts; however, comparable Bayesian models with random slopes revealed the same pattern of effects (Appendix B, Table 1).



**Figure 2.2. Order memory performance.** (a) Model estimates (i.e., estimated marginal means) of temporal order accuracy during predictable (“pred”) versus variable (“var”) events, adjusted for several confound variables (i.e., block, stimulus category, aisle response RTs). (b) Visualization of raw order memory accuracy across blocks, showing the qualitative pattern that enhanced memory for predictable events emerges gradually across blocks. Error bars in (a) indicate 95% confidence intervals from fitted model, and error bars in (b) indicate within-subjects standard errors. \* $p < .05$

Replicating our previous work (Gasser & Davachi, 2023; Chapter 1), we found a significant effect of condition on order accuracy ( $b = 0.033$ ,  $SE = 0.013$ ,  $CI = [0.008, 0.059]$ ,  $p = .010$ ), such that order memory was more accurate for items encoded in the predictable compared to the variable store

(Figure 2.2a). There were also significant effects of block ( $b = 0.016$ ,  $SE = 0.002$ ,  $CI = [0.012, 0.020]$ ,  $p < .001$ ) and stimulus category ( $b = -0.050$ ,  $SE = 0.013$ ,  $CI = [-0.076, -0.024]$ ,  $p < .001$ ), such that order accuracy was better during later blocks and for events with food (vs. animal) items. The effect of aisle response RTs within an event on subsequent order memory was not significant ( $b = 0.005$ ,  $SE = 0.031$ ,  $CI = [-0.054, 0.068]$ ,  $p = .822$ ). Although qualitative examination of order memory performance suggests that the performance enhancement for predictable versus variable events emerged gradually across blocks (Figure 2.2b), the interaction between condition and block did not reach significance in our model ( $b = 0.005$ ,  $SE = 0.004$ ,  $CI = [-0.002, -0.013]$ ,  $p = .164$ ).

In this first model, we quantified order memory test performance by calculating the proportion of items selected in the correct ordinal position (e.g., first, second). However, this measure may not account for situations where participants accurately remembered sub-sequences of objects, but selected them in the wrong ordinal positions. To address this limitation, we also computed Levenshtein distance as a second order memory metric, which quantifies the minimum number of “edits” (i.e., insertions, deletions, or substitutions) needed to transform the participants’ recalled object sequence into the correctly ordered sequence (Levenshtein, 1966). When using this measure as our outcome variable, the effect of condition on order memory performance remained significant ( $b = -0.157$ ,  $SE = 0.066$ ,  $CI = [-0.288, -0.027]$ ,  $p = 0.018$ ), as did the effects of block ( $b = -0.094$ ,  $SE = 0.010$ ,  $CI = [-0.114, -0.075]$ ,  $p < .001$ ) and stimulus category ( $b = 0.285$ ,  $SE = 0.067$ ,  $CI = [0.154, 0.417]$ ,  $p < .001$ ). The effect of aisle response RTs was again non-significant ( $b = -0.018$ ,  $SE = 0.160$ ,  $CI = [-0.332, 0.296]$ ,  $p = 0.912$ ), although the interaction between block and condition was marginal ( $b = -0.034$ ,  $SE = 0.019$ ,  $CI = [-0.072, 0.004]$ ,  $p = 0.078$ ). Taken together, these results affirm our previous findings that the execution of a well-learned action sequence enhances temporal order memory for concurrent episodic information.

### B.3.4 Object recognition performance

At the end of each experiment day, participants completed bouts of an object recognition memory test (Figure 2.1d). During this test, they had to discriminate between objects that had been presented during their errands and perceptually similar lures. This recognition test was structured such that half of

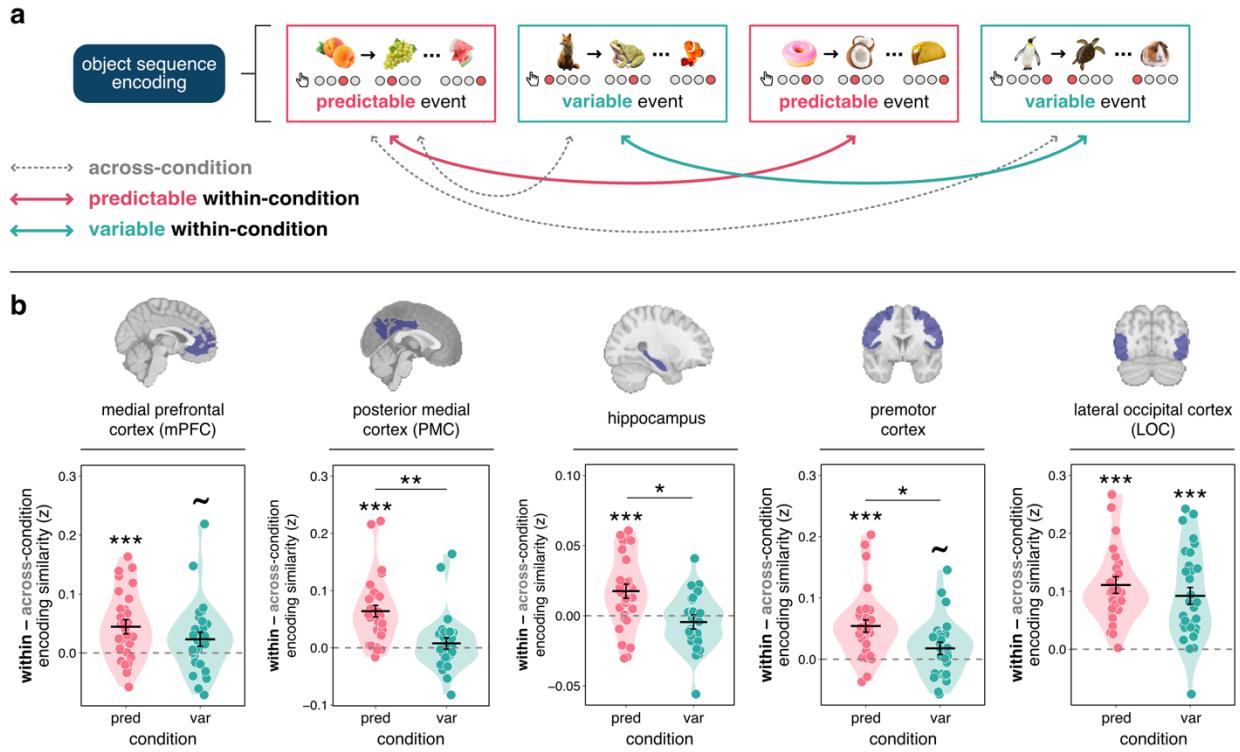
the tested objects were drawn from errands that occurred on the same day as the test (“immediate” memory), and half were drawn from errands that occurred on the previous day (“delay” memory). To explore whether execution of the predictable action sequence affected object recognition, we ran another series of regression models analogous to those used to explore order memory, in which the proportion of objects successfully recognized (or whose similar lures were correctly rejected) within an event was predicted by condition, block number, stimulus category, and the mean (log-transformed) aisle response RT within the event. Each model included an interaction term between condition and block. Separate models were run for events that were included as part of the immediate versus delay tests. Both models included only random intercepts due to convergence issues when including random slopes, but results were the same when using Bayesian models with the full set of random effects (Appendix B, Table 2).

We found that recognition memory performance was not significantly associated with condition at either the immediate ( $b = -0.004$ ,  $SE = 0.012$ ,  $CI = [-0.028, 0.020]$ ,  $p = .728$ ) or delay ( $b = 0.025$ ,  $SE = 0.014$ ,  $CI = [-0.002, 0.053]$ ,  $p = .074$ ) timepoints. There were also no effects of encoding block (immediate:  $b = 0.002$ ,  $SE = 0.002$ ,  $CI = [-0.001, 0.006]$ ,  $p = .234$ ; delay:  $b = 0.003$ ,  $SE = 0.002$ ,  $CI = [-0.001, 0.008]$ ,  $p = .123$ ) or aisle response RTs (immediate:  $b = 0.016$ ,  $SE = 0.028$ ,  $CI = [-0.040, 0.071]$ ,  $p = 0.579$ ; delay:  $b = 0.032$ ,  $SE = 0.033$ ,  $CI = [-0.032, 0.097]$ ,  $p = .330$ ), nor was there a significant interaction between condition and block (immediate:  $b = 0.004$ ,  $SE = 0.004$ ,  $CI = [-0.003, 0.011]$ ,  $p = .230$ ; delay:  $b = -0.005$ ,  $SE = 0.004$ ,  $CI = [-0.014, 0.003]$ ,  $p = .191$ ). However, at both test timepoints, recognition performance was better for food versus animal items (immediate:  $b = -0.097$ ,  $SE = 0.012$ ,  $CI = [-0.122, -0.073]$ ,  $p < .001$ , delay:  $b = -0.055$ ,  $SE = 0.014$ ,  $CI = [-0.083, -0.027]$ ,  $p < .001$ ). These results corroborate our previous behavioral findings that the effects of specific motor sequence knowledge are selective to temporal order memory (Gasser & Davachi, 2023; Chapter 1).

### B.3.5 Neural representation of the predictable action scaffold

Thus far, we have shown that when participants execute a predictable action sequence during a novel episodic event, they have enhanced order memory for items within that event. This suggests that the predictable action sequence scaffolds memory for concurrent object sequences, above and beyond any

memory benefits conferred by less specific kinds of prior knowledge (i.e., about the semantic category of objects). In the next set of analyses, we explored how this action scaffold was represented in the brain.

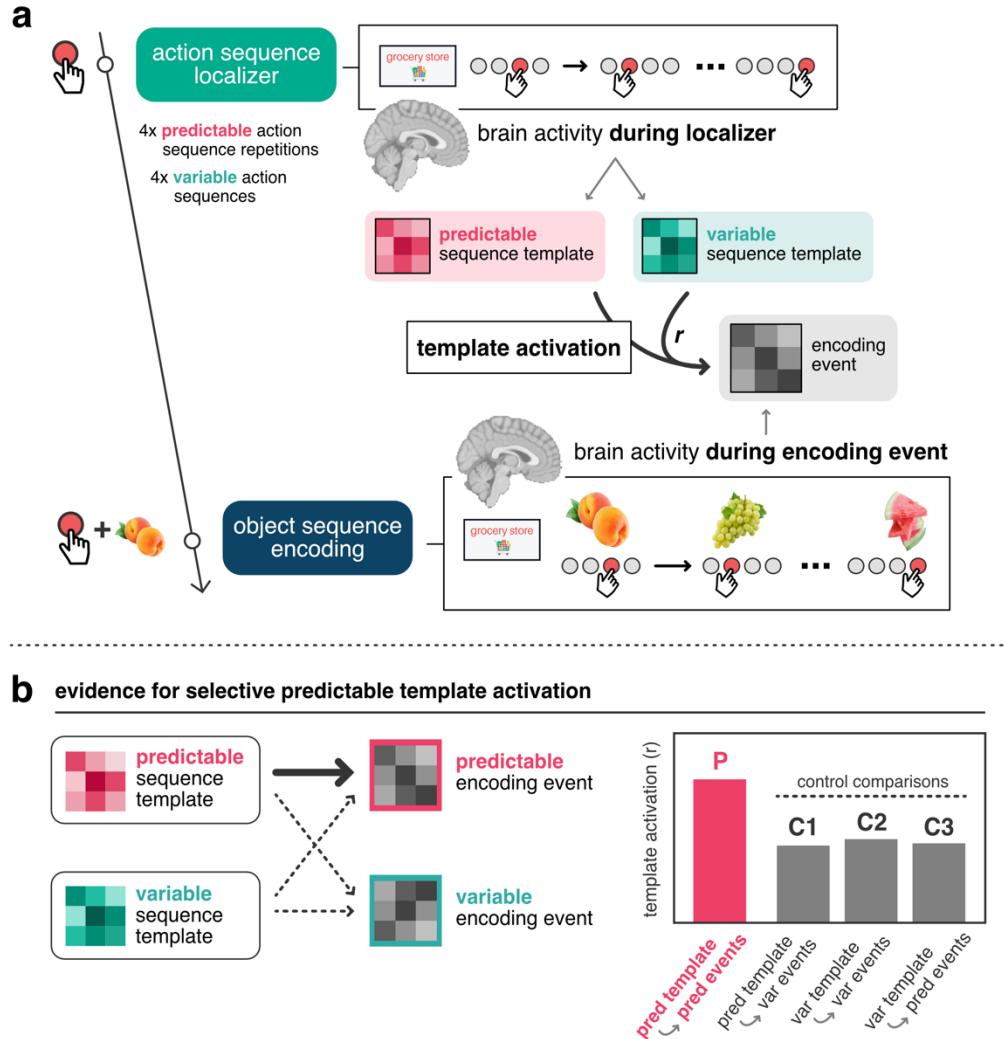


**Figure 2.3. Encoding pattern similarity across events.** (a) Illustration of the analysis procedure. Within each ROI, a single brain activity pattern was extracted for each encoding event. We then computed correlations across event patterns to derive three similarity metrics: the average similarity of predictable events to each other (“predictable within-condition”); the average similarity of variable events to each other (“variable within-condition”); and the average similarity of event pairs from different conditions (“across-condition”). (b) Within-condition (minus across-condition) similarity for predictable versus variable events. Values that exceed zero (the dotted line) indicate that event patterns within a given condition were more similar to each other than to events in the other condition. Significance stars above each cloud of datapoints indicate whether within-condition similarity exceeds across-condition similarity. Significance stars above lines that bridge across datapoint clouds indicate whether the subtraction between within- and across-condition similarity was different for predictable versus variable events. In all plots, error bars indicate within-subjects standard errors, and all p-values are FDR-corrected for multiple comparisons across ROIs. ~ $p < .1$ , \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

To begin, we reasoned that the consistent motor structure of the predictable store should be reflected in the similarity of brain activity patterns *across events*. Specifically, if a region represents this action scaffold, activity patterns from predictable encoding events should be more similar to each other

than they are to patterns from variable events (which lack the learned action sequence). To test this hypothesis, we computed the similarity between each event's encoding pattern and all other encoding patterns from the same condition (*within-condition* similarity) versus all patterns from the other condition (*across-condition* similarity); Figure 2.3a. For predictable events, we found that within-condition similarity was significantly greater than across-condition similarity in all of our targeted regions of interest (Figure 2.3b), including mPFC ( $t(29) = 4.224$ ,  $CI = [0.023, 0.066]$ ,  $p_{FDR} < .001$ ,  $d = 0.771$ ), PMC ( $t(29) = 3.927$ ,  $CI = [0.027, 0.086]$ ,  $p_{FDR} < .001$ ,  $d = 1.106$ ), premotor cortex ( $t(29) = 5.279$ ,  $CI = [0.033, 0.075]$ ,  $p_{FDR} < .001$ ,  $d = 0.964$ ), hippocampus ( $t(29) = 3.684$ ,  $CI = [0.008, 0.027]$ ,  $p_{FDR} < .001$ ,  $d = 0.673$ ), and LOC ( $t(29) = 9.051$ ,  $CI = [0.086, 0.136]$ ,  $p_{FDR} < .001$ ,  $d = 1.653$ ).

Given that the variable context also contains some learned structure, in that participants are familiar with the store identity and its associated object category, we also asked whether variable events evoked a selective neural representation. Indeed, within-condition similarity for variable events was significantly greater than across-condition similarity in LOC ( $t(29) = 6.458$ ,  $CI = [0.063, 0.121]$ ,  $p_{FDR} < .001$ ,  $d = 1.179$ ), with this comparison reaching marginal significance in both mPFC ( $t(29) = 2.152$ ,  $CI = [0.001, 0.045]$ ,  $p_{FDR} = .068$ ,  $d = 0.398$ ) and premotor cortex ( $t(29) = 2.143$ ,  $CI = [0.001, 0.035]$ ,  $p_{FDR} = .068$ ,  $d = 0.391$ ). No such effect was observed in PMC ( $t(29) = 0.865$ ,  $CI = [-0.011, 0.026]$ ,  $p_{FDR} = .394$ ,  $d = 0.158$ ) or the hippocampus ( $t(29) = -1.286$ ,  $CI = [-0.012, 0.003]$ ,  $p_{FDR} = .261$ ,  $d = 0.235$ ). Finally, we compared the difference between within-condition and across-condition similarity for predictable versus variable events. This analysis revealed that activity patterns were more consistent across predictable than across variable events in PMC ( $t(29) = 3.927$ ,  $CI = [0.027, 0.086]$ ,  $p_{FDR} = .002$ ,  $d = 1.052$ ), premotor cortex ( $t(29) = 2.575$ ,  $CI = [0.008, 0.065]$ ,  $p_{FDR} = .026$ ,  $d = 0.716$ ), and the hippocampus ( $t(29) = 3.119$ ,  $CI = [0.008, 0.036]$ ,  $p_{FDR} = .010$ ,  $d = 0.972$ ), but not in mPFC ( $t(29) = 1.253$ ,  $CI = [-0.013, 0.056]$ ,  $p_{FDR} = .275$ ,  $d = 0.365$ ) or LOC ( $t(29) = 0.914$ ,  $CI = [-0.023, 0.061]$ ,  $p_{FDR} = .369$ ,  $d = 0.257$ ). Taken together, these results suggest that predictable encoding events are associated with a consistent and unique neural signature across default mode and sensorimotor regions, exceeding the extent to which the variable context is represented in a subset of targeted ROIs.



**Figure 2.4. Illustration of template activation analysis.** (a) First, we extracted “sequence templates” for the predictable and variable conditions, which reflect the average activity patterns evoked while participants executed predictable or variable action sequences during the localizer task. Then, we computed the similarity between these template representations and patterns evoked during each encoding event. (b) If a region represents the predictable action sequence, we should see stronger similarity between the predictable template and predictable encoding events (“P”) than any control template activation measure (C1, C2, and C3). See Methods for more details on the interpretation of each control.

While the previous analysis tells us that predictable events tend to elicit more similar or stable activity patterns, we can only infer that this effect is driven by the activation of the action scaffold. To directly test for the presence of this scaffold representation, we used neural “templates” derived from the independent localizer task that took place prior to encoding (Figure 2.1a), during which participants executed action sequences (both predictable and variable) without seeing objects. Specifically, we

extracted templates corresponding to the average patterns of brain activity evoked while participants performed predictable or variable motor sequences (Figure 2.4a; also see *Methods*). We then computed the similarity between these templates and brain activity evoked during the subsequent object sequence encoding task. To the extent that a given brain region represents the predictable action scaffold, we should see high similarity between the predictable sequence template in that region and activity patterns from predictable encoding events. This critical measure of “predictable template activation” (P) was validated by comparing it to three control similarity measures (C1, C2, and C3), which broadly ensured that evidence for the predictable template was selective to the predictable encoding context. See *Methods* for details about the computation and interpretation of each individual control measure.

In mPFC, evidence for the predictable sequence template significantly exceeded all three control comparisons ( $P > C1: t(29) = 2.833, CI = [0.013, 0.083], p_{FDR} = .014, d = 0.345$ ;  $P > C2: t(29) = 3.144, CI = [0.026, 0.122], p_{FDR} = .006, d = 0.586$ ;  $P > C3: t(29) = 2.797, CI = [0.019, 0.122], p_{FDR} = .023, d = 0.521$ ). Similar results were found in PMC, although the third control comparison reached only marginal significance ( $P > C1: t(29) = 6.208, p_{FDR} < .001, CI = [0.047, 0.093], d = 0.475$ ;  $P > C2: t(29) = 3.689, CI = [0.043, 0.149], p_{FDR} = .002, d = 0.682$ ;  $P > C3: t(29) = 1.984, CI = [-0.002, 0.118], p_{FDR} = .095, d = 0.373$ ). The predictable sequence template was also strongly activated in premotor cortex ( $P > C1: t(29) = 4.816, CI = [0.044, 0.108], p_{FDR} < .001, d = 0.368$ ;  $P > C2: t(29) = 3.927, CI = [0.057, 0.182], p_{FDR} = .002, d = 0.621$ ;  $P > C3: t(29) = 2.852, CI = [0.027, 0.165], p_{FDR} = .023, d = 0.470$ ). See Figure 2.5a for these effects.

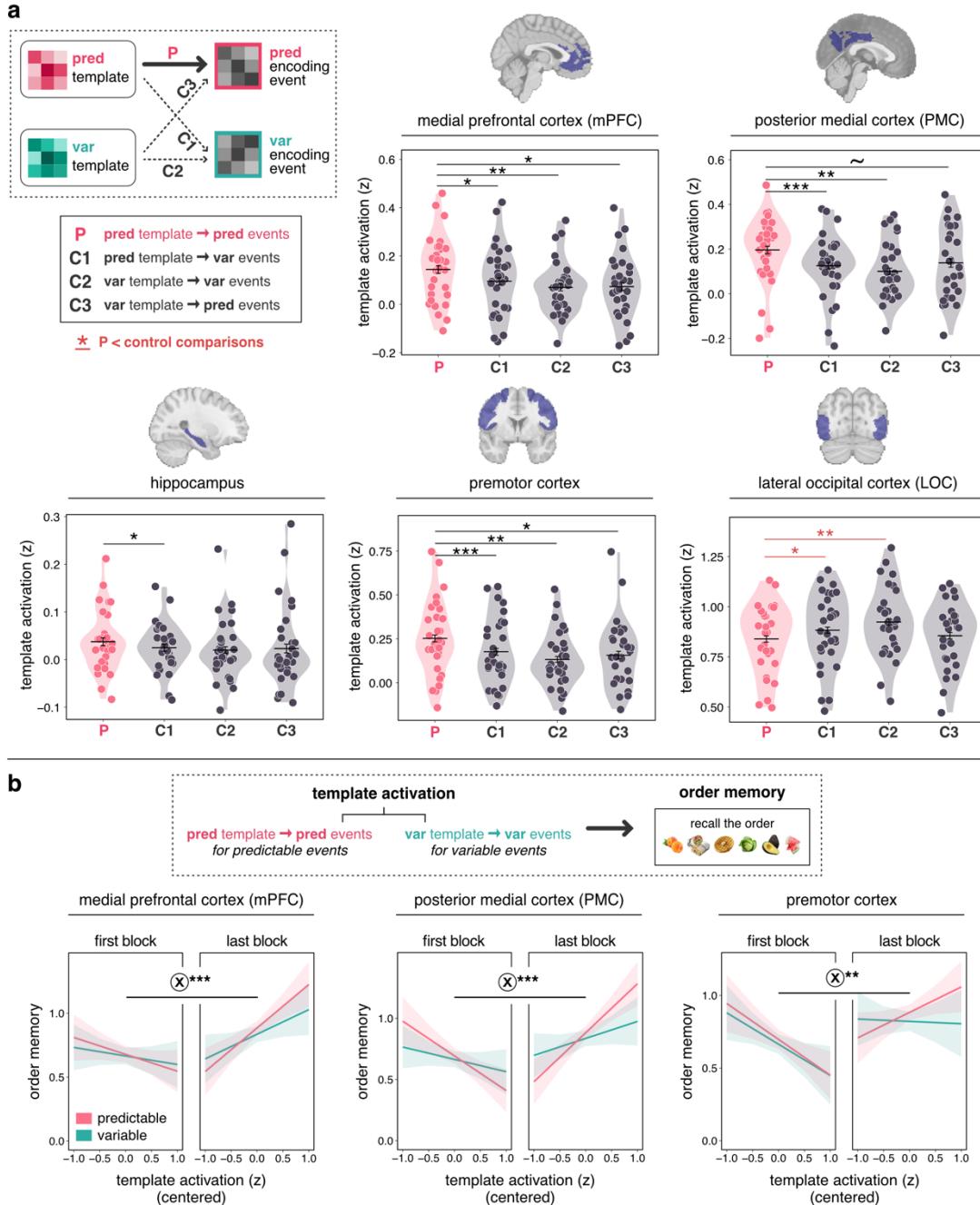
Next, we turned to the hippocampus. Although the hippocampus contained distinctive representations of the predictable context during encoding (as reported in previous set of analyses; Figure 2.3b), there was not compelling evidence that this representation contained the specific predictable action sequence. In the hippocampus, only one of our three control comparisons reached significance ( $P > C1: t(29) = 2.245, CI = [0.001, 0.024], p_{FDR} = .032, d = 0.195$ ;  $P > C2: t(29) = 1.386, CI = [-0.008, 0.043], p_{FDR} = .176, d = 0.270$ ;  $P > C3: t(29) = 1.009, CI = [-0.015, 0.043], p_{FDR} = .402, d = 0.186$ ); Figure 2.5a. Selective representation of the predictable action scaffold was also not observed in LOC. In fact, two out

of the three control comparisons in this region reached significance in the *opposite* direction as what would be expected by predictable template activation, in that similarity between the predictable sequence template and predictable encoding events (P) was *lower* than similarity between the predictable template and variable events (C1), as well as between the variable template and variable events (C2); P > C1:  $t(29) = -2.708$ ,  $CI = [-0.077, -0.011]$ ,  $p_{FDR} = .014$ ,  $d = 0.244$ ; P > C2:  $t(29) = -3.026$ ,  $CI = [-0.142, -0.028]$ ,  $p_{FDR} = .006$ ,  $d = 0.490$ ; P > C3:  $t(29) = -0.785$ ,  $CI = [-0.059, 0.026]$ ,  $p_{FDR} = .439$ ,  $d = 0.010$  (Figure 2.5a).

Although it is not immediately clear how to interpret these opposite effects, we can nevertheless infer that LOC encoding patterns do not reflect predictable motor behavior.

Next, to characterize the spatial and temporal specificity of template activation, we performed two exploratory follow-up analyses. First, we ran a whole-brain searchlight analysis to see where the predictable action scaffold was represented beyond our targeted ROIs. Briefly, this searchlight identified regions where predictable template activation exceeded all three control measures. Evidence for predictable template activation was seen in several patches of cortex overlapping with our *a priori* ROIs (Appendix B, Figure 1), including mPFC (at a relaxed statistical threshold), PMC (i.e., precuneus and posterior cingulate), and premotor cortex. Clusters were also observed in several other sensorimotor regions (e.g., SMA, pre-SMA, M1), as well as in lateral prefrontal cortex and lateral parietal cortex. Second, instead of looking for template activation that was sustained throughout the duration of the event, we considered the possibility that the strength of this representation may peak at specific moments in time, e.g., at the beginning of an event, when participants must prepare to execute the learned action sequence. To explore this idea, we computed template activation separately during each individual sequence position. In mPFC, PMC, and premotor cortex, predictable template activation was relatively consistent throughout the event, dropping off only at the final sequence position (Appendix B, Figure 2).

***Relationship between template activation and order memory.*** We hypothesized that the representation of the predictable action sequence functions as a scaffold for incoming information, allowing novel item sequences to become rapidly integrated with existing motor knowledge. This framework suggests that stronger neural representation of the predictable action scaffold should be



**Figure 2.5. Template activation.** (a) Predictable template activation ( $P$ , i.e., similarity between the predictable sequence template and predictable encoding event patterns) versus control measures ( $C1$ ,  $C2$ , and  $C3$ ) in all ROIs. A region is said to represent the predictable action scaffold if  $P$  exceeds all controls. Red significance stars signify comparisons where a control measure exceeds  $P$ . Error bars show within-subjects standard errors. (b) Relationship between template activation and order memory for predictable and variable events. Plots show fitted regression lines and 95% confidence intervals from multilevel models where order accuracy for each event was predicted by template activation, condition, and block number. Interaction symbols ( $\times$ ) indicate the significance of the template activation by block interaction term. Note that block was modeled as a continuous variable; we include separate panels for the first and last blocks for visualization purposes. All p-values are FDR-corrected for multiple comparisons across ROIs.  $\sim p < .1$ ,  $*p < .05$ ,  $**p < .01$ ,  $***p < .001$

linked to better order memory for items encoded during predictable events. To explore this idea, a series of multilevel regression models were constructed in which order accuracy for each event (i.e., the proportion of items in the event recalled in the correct temporal order) was predicted by activation of the relevant template (i.e., the predictable template for predictable events and the variable template for variable events; Figure 2.5b). These models also included predictors for condition (predictable vs. variable), as well as for block number (given tentative evidence that the order memory effect evolved across time; Figure 2.2b), as well as all possible interaction terms between the three predictors. Separate models were run within mPFC, PMC, and premotor cortex, all of which showed evidence for predictable template activation (Figure 2.5a).

In mPFC, there was no main effect of template activation on order memory ( $b = 0.084$ ,  $SE = 0.038$ ,  $CI = [0.002, 0.165]$ ,  $p_{FDR} = .133$ ), but there was a significant interaction between template activation and block ( $b = 0.033$ ,  $SE = 0.009$ ,  $CI = [0.015, 0.052]$ ,  $p_{FDR} < .001$ ), such that template activation became more positively associated with better order memory in later task blocks (Figure 2.5b). However, there was no interaction between template activation in mPFC and condition ( $b = 0.041$ ,  $SE = 0.069$ ,  $CI = [-0.095, 0.177]$ ,  $p_{FDR} = .572$ ), indicating that the beneficial effect of template activation on temporal order memory was not specific to the predictable context. Both PMC and premotor cortex showed a similar pattern of effects (Figure 2.5b). That is, although there was no main effect of template activation on order memory in either region (PMC:  $b = 0.039$ ,  $SE = 0.041$ ,  $CI = [-0.046, 0.124]$ ,  $p_{FDR} = .355$ ; premotor cortex:  $b = -0.076$ ,  $SE = 0.045$ ,  $CI = [-0.169, 0.017]$ ,  $p_{FDR} = .160$ ), there were significant interactions between template activation and block (PMC:  $b = 0.042$ ,  $SE = 0.009$ ,  $CI = [0.024, 0.060]$ ,  $p_{FDR} < .001$ ; premotor cortex:  $b = 0.028$ ,  $SE = 0.009$ ,  $CI = [0.010, 0.047]$ ,  $p_{FDR} = .003$ ). As with mPFC, neither PMC ( $b = 0.040$ ,  $SE = 0.070$ ,  $CI = [-0.098, 0.178]$ ,  $p_{FDR} = .572$ ) nor premotor cortex ( $b = 0.080$ ,  $SE = 0.073$ ,  $CI = [-0.065, 0.225]$ ,  $p_{FDR} = .572$ ) showed an interaction between condition and template activation. However, in PMC, the three-way interaction between condition, template activation, and block reached marginal significance ( $b = 0.040$ ,  $SE = 0.018$ ,  $CI = [0.006, 0.075]$ ,  $p_{FDR} = .066$ ), suggesting that the extent to which the association between template activation and order memory grew more positive

across blocks was slightly greater for predictable versus variable events. For full model output, see Appendix B, Table 3.

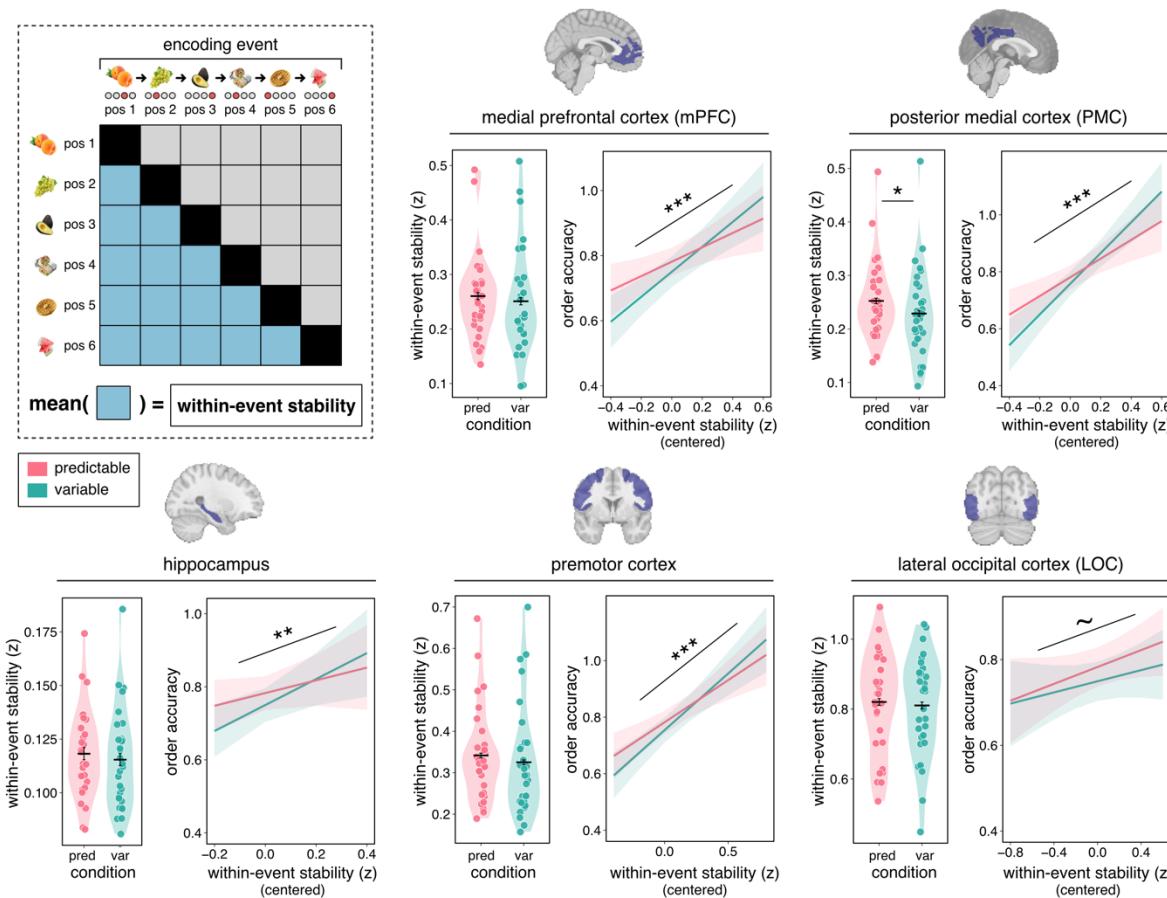
Ultimately, these results indicate that the extent to which participants represented the learned motor structure during predictable events bolstered subsequent order memory, specifically during later blocks of the experiment. This finding lends credence to our idea that this action sequence functions as a scaffold for concurrent episodic information. However, we also observed that stronger representation of the *variable* template (during variable encoding events) was similarly beneficial for order memory. Even though the variable template contains no consistent motor sequence, it may still function as an abstract representation of the variable context, holding information about the identity of the variable store, its perceptual characteristics, and the general structure of variable events (e.g., having a sequence of six motor actions). As such, it may be that irrespective of whether a given context is associated with a predictable motor sequence, the representation of that context during encoding is associated with improved subsequent order memory — particularly as participants gain experience with our task.

### B.3.6 Stability of brain activity within encoding events

So far, we have established that predictable encoding events evoke a consistent neural activity pattern that represents the learned action sequence in mPFC, PMC, and premotor cortex. Our results also suggest that the strength with which context-specific neural representations in these regions are activated during encoding supports order memory — specifically during later blocks of our task. However, these results prompt the question of how exactly the predictable action scaffold is exerting its effects on memory. As described in the *Introduction*, previous research suggests that one mechanism supporting temporal memory may be the stability of neural activity across time (DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014). Motivated by these findings, we next examined how the execution of a predictable action sequence modulates the stability of neural activity *within an event*, as well as how such stability relates to memory. We quantified “within-event stability” by obtaining brain activity patterns corresponding to each individual sequence position during encoding (see *Methods*; Figure 2.6), and then computing the average similarity between those patterns within each event. Then, we compared the strength of this stability

metric between predictable versus variable events.

Interestingly, in PMC ( $t(29) = 3.154, CI = [0.008, 0.039], p_{FDR} = .019, d = 0.285$ ), within-event stability was significantly greater for predictable relative to variable events, suggesting the presence of the predictable action scaffold enhances the stability of concurrent neural activity (Figure 2.6). In contrast, within-event stability did not differ by condition in mPFC ( $t(29) = 1.027, CI = [-0.009, 0.028], p_{FDR} = .497, d = 0.100$ ), premotor cortex ( $t(29) = 2.061, CI = [0.0001, 0.033], p_{FDR} = .121, d = 0.117$ ), the hippocampus ( $t(29) = 0.687, CI = [-0.005, 0.011], p_{FDR} = .497, d = 0.124$ ), or LOC ( $t(29) = 0.739, CI = [-0.018, 0.038], p_{FDR} = .497, d = 0.070$ ); Figure 2.6.



**Figure 2.6. Within-event stability.** The left plots in each ROI panel show the difference in within-event stability during predictable versus variable events. In these plots, error bars represent within-subject standard errors. The right plots in each panel show the relationship between order accuracy and within-event stability. These plots depict fitted regression lines and 95% confidence intervals from multilevel models where order accuracy for each event was predicted by within-event stability, condition, and block number. All p-values are FDR-corrected for multiple comparisons across ROIs.  $\sim p < .1$ ,  $*p < .05$ ,  $**p < .01$ ,  $***p < .001$

Next, we examined the relationship between within-event stability during encoding and subsequent order memory. Analogous to our approach with template activation, we implemented a series of regression models in which order memory accuracy (i.e., the proportion of items selected in the correct order within an event) was predicted by stability, condition, block number, and all two- and three-way interactions between those terms. In general, within-event stability was robustly associated with better order memory, both for predictable and variable events (Figure 2.6). The main effect of stability on memory was significant in mPFC ( $b = 0.302$ ,  $SE = 0.066$ ,  $CI = [0.167, 0.437]$ ,  $p_{FDR} < .001$ ), PMC ( $b = 0.434$ ,  $SE = 0.073$ ,  $CI = [0.284, 0.584]$ ,  $p_{FDR} < .001$ ), hippocampus ( $b = 0.265$ ,  $SE = 0.098$ ,  $CI = [0.072, 0.485]$ ,  $p_{FDR} = .009$ ), and premotor cortex ( $b = 0.349$ ,  $SE = 0.059$ ,  $CI = [0.226, 0.472]$ ,  $p_{FDR} < .001$ ). Stability in LOC, conversely, was only marginally associated with subsequent order memory ( $b = 0.082$ ,  $SE = 0.045$ ,  $CI = [-0.006, 0.169]$ ,  $p_{FDR} = .067$ ). In all ROIs, the relationship between stability and memory performance did not significantly interact with condition or block number. For the full model output, see Appendix B, Table 4, as well as Table 5 for supplementary Bayesian models.

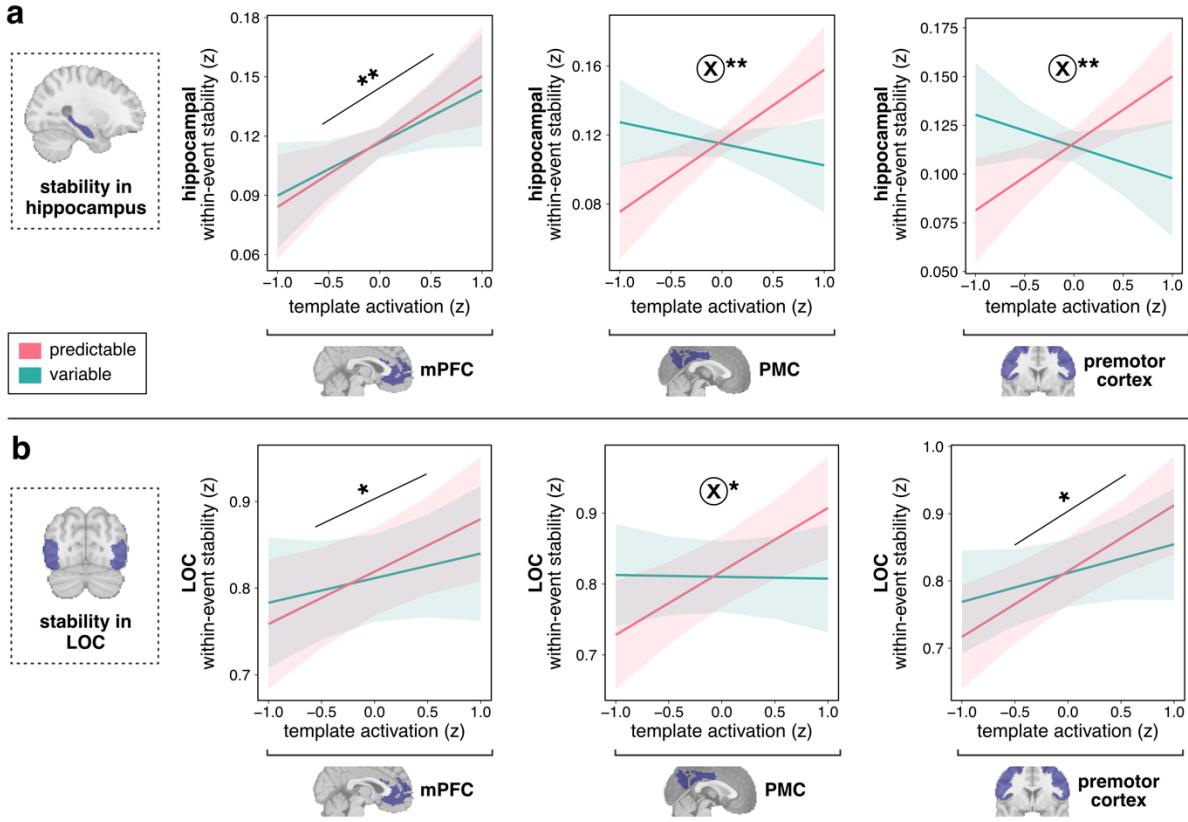
Taken together, these results indicate that the stability of neural activity supports the formation of temporally structured event memories. Specifically, in all of our ROIs (except for LOC, where the effect was marginal), greater stability during encoding was significantly associated with enhanced subsequent order memory. This relationship was similar in both the predictable and variable encoding contexts, suggesting that stability serves as a general mechanism of temporal memory encoding that is beneficial irrespective of participants' prior motor sequence knowledge. That being said, the fact that stability in PMC was generally stronger during predictable events indirectly suggests that having specific motor sequence knowledge might upregulate the degree of neural stability during encoding, which in turn bolsters temporal memory.

### B.3.7 Connecting template activation and within-event stability

The ultimate goal of this study was to understand how specific motor sequence knowledge (i.e., the predictable action scaffold) modulates mechanisms of temporal order encoding. Thus far, we have shown that predictable encoding events evoke neural representations of that action scaffold, and that the

stability of brain activity throughout encoding is beneficial for subsequent order memory. For the final set of analyses, we sought to connect these two measures together, asking how representation of the predictable action sequence modulates within-event stability. To that end, we asked whether template activation — seen to be robust and significant in mPFC, PMC, and premotor cortex — was related to within-event stability in the hippocampus and LOC. We focused in particular on the hippocampus and LOC because we hypothesized that stability in these regions would most directly support the encoding of novel object sequences, as found in or previous work (DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014). Furthermore, the hippocampus and LOC are known to play critical roles in supporting memory for temporal sequences (Eichenbaum, 2017; Ranganath & Hsieh, 2016; Tubridy & Davachi, 2011) and for visual objects (Grill-Spector et al., 2001), respectively. For this set of analyses, we ran a series of linear regression models in which within-event stability was predicted by template activation, condition, and block (plus all interactions between these terms). For predictable events, template activation refers to similarity between the event pattern and the predictable template; for variable events, template activation is similarity between the event pattern and the variable template. Note that to allow for model convergence, all models reported in this section include only random intercepts. However, unless otherwise noted, all effects of interest survived in supplemental Bayesian models that include the full set of random effects.

First, we looked at how template activation in mPFC, PMC, and premotor cortex was associated with stability in the hippocampus (Figure 2.7a). In mPFC, greater template activation was associated with greater hippocampal stability in both encoding conditions ( $b = 0.030$ ,  $SE = 0.009$ ,  $CI = [0.011, 0.048]$ ,  $p_{FDR} = .005$ ). Interestingly, in both PMC and premotor cortex, this positive association between template activation and stability was selective to the predictable condition. That is, while neither region showed a main effect of template activation on stability (PMC:  $b = 0.014$ ,  $SE = 0.009$ ,  $CI = [-0.004, 0.033]$ ,  $p_{FDR} = .189$ ; premotor cortex:  $b = 0.009$ ,  $SE = 0.002$ ,  $CI = [-0.011, 0.029]$ ,  $p_{FDR} = .370$ ), both showed significant interactions between template activation and condition (PMC:  $b = 0.054$ ,  $SE = 0.017$ ,  $CI = [0.020, 0.088]$ ,  $p_{FDR} = .004$ ; premotor cortex:  $b = 0.051$ ,  $SE = 0.017$ ,  $CI = [0.018, 0.083]$ ,  $p_{FDR} = .004$ ). A follow-up



**Figure 7. Association between template activation and within-event stability.** (a) Relationship between within-event stability in the hippocampus and template activation in mPFC (left), PMC (middle), and premotor cortex (right). (b) Relationship between within-event stability in LOC and template activation in the same three regions. Note that template activation refers to similarity between an event pattern and the predictable sequence template for predictable events, and between an event pattern and the variable sequence template for variable events. Plots show fitted regression lines and 95% confidence intervals from multilevel models where stability for each event was predicted by template activation, condition, and block number. Interaction symbols ( $\otimes$ ) indicate the significance of the template activation by condition interaction term. All p-values are FDR-corrected for multiple comparisons across ROIs.  $\sim p < .1$ ,  $*p < .05$ ,  $**p < .01$ ,  $***p < .001$

analysis of simple slopes (assessing the average effect of template activation on stability, separately for each condition) revealed that these interactions were driven by significant positive relationships in the predictable condition (PMC:  $b = 0.041$ ,  $SE = 0.013$ ,  $p_{FDR} = .004$ ; premotor cortex:  $b = 0.034$ ,  $SE = 0.012$ ,  $p_{FDR} = 0.008$ ), and null effects in the variable condition (PMC:  $b = -0.013$ ,  $SE = 0.013$ ,  $p_{FDR} = .147$ ; premotor cortex:  $b = -0.016$ ,  $SE = 0.014$ ,  $p_{FDR} = .330$ ). Of note, the relationship between hippocampal stability and template activation (in mPFC, PMC, and premotor cortex) did not significantly change

across blocks. See Appendix B, Table 6 for full model output and Table 7 for supplementary Bayesian model results. In sum, stronger representation of the predictable sequence template in both PMC and premotor cortex enhanced hippocampal stability selectively for predictable events.

We next performed an analogous set of analyses looking at neural stability in LOC (Figure 2.7b). Template activation in mPFC was associated with greater LOC stability across both conditions ( $b = 0.045$ ,  $SE = 0.021$ ,  $CI = [0.004, 0.085]$ ,  $p_{FDR} = .037$ ), although this effect was not statistically reliable in a supplemental Bayesian model (Appendix B, Table 9). The relationship between template activation in mPFC and LOC stability also did not interact with condition ( $b = 0.032$ ,  $SE = 0.038$ ,  $CI = [-0.043, 0.107]$ ,  $p_{FDR} = .398$ ). Effects were similar for template activation in premotor cortex, such that there was a significant main effect of template activation on LOC stability ( $b = 0.070$ ,  $SE = 0.024$ ,  $CI = [0.023, 0.117]$ ,  $p_{FDR} = .011$ ) that did not vary by condition ( $b = 0.055$ ,  $SE = 0.036$ ,  $CI = [-0.015, 0.125]$ ,  $p_{FDR} = .189$ ). In PMC, however, template activation was selectively associated with LOC stability for predictable but not variable events — mirroring the condition-specific relationship between PMC template activation and hippocampal stability. Specifically, this model revealed a significant main effect of template activation ( $b = 0.044$ ,  $SE = 0.021$ ,  $CI = [0.003, 0.085]$ ,  $p_{FDR} = .037$ ), as well as a significant interaction between template activation and condition ( $b = 0.092$ ,  $SE = 0.037$ ,  $CI = [0.019, 0.165]$ ,  $p_{FDR} = .039$ ). A follow-up simple slopes analysis revealed that this interaction was driven by a positive association between template activation in PMC and stability in LOC during predictable ( $b = 0.090$ ,  $SE = 0.028$ ,  $p_{FDR} = .002$ ) but not variable ( $b = -0.003$ ,  $SE = 0.028$ ,  $p_{FDR} = .927$ ) events. See Appendix B, Table 8 for full model output and Table 9 for supplementary Bayesian models.

For this set of analyses, we focused on stability in the hippocampus and LOC based on the reasoning that these regions are most directly linked by prior work to the encoding of novel objects. However, for completeness, we also checked whether our measures of template activation and stability were correlated within the same ROI. In theory, greater neural stability across time in a given region might increase our chances of detecting template activation in that region — given that our main template activation analysis implicitly searches for representations that are maintained throughout the

duration of the event (see *Methods*). Indeed, in all regions showing significant predictable template activation (mPFC, PMC, and premotor cortex), we found that template activation was positively associated with stability in those same cortical targets across both conditions. Although there was some evidence that the association between these measures was stronger for predictable versus variable events in PMC and premotor cortex, these effects were not consistent across our two modeling approaches (i.e., a random intercept-only Frequentist model and a Bayesian model with random slopes). See Appendix B, Tables 10 and 11 for these supplemental results.

In sum, this final set of analyses reveals a putative mechanism for how the predictable action scaffold exerts its effects on temporal order memory. In PMC, greater activation of the predictable sequence template selectively stabilized activity in the hippocampus and LOC during predictable encoding events. Greater stability in the hippocampus, in turn, was related to improved order memory performance. In contrast, template activation in mPFC stabilized hippocampal activity similarly across conditions, perhaps indicating that specific motor sequence knowledge does not modulate how representations in this region impact episodic memory formation. Premotor cortex showed a more heterogenous pattern of effects, with template activation relating to hippocampal stability selectively for predictable events, but to stability in LOC across both conditions.

## B.4 Discussion

In this study, we explored how predictable sequence knowledge scaffolds temporal order memory for concurrent object sequences. First, we showed that executing well-learned motor behavior during encoding events elicited stable and distinctive patterns of brain activity in both default mode network and sensorimotor regions. In contrast, distinct neural representations of the variable encoding context were generally weaker or absent altogether. We further found that the content of this predictable neural signature reflected learned motor sequence knowledge in medial prefrontal cortex (mPFC), posterior medial cortex (PMC), and premotor cortex. Importantly, the strength of these knowledge representations was associated with enhanced temporal order memory, specifically during the later encoding blocks of

our experiment. We also found that temporal order memory was linked to the stability of neural activity across time in both the predictable and variable contexts, and that in PMC, within-event stability was significantly greater for predictable versus variable events. Finally, we found that stronger representations of the predictable action scaffold in PMC and premotor cortex were associated with within-event stability in the hippocampus and LOC, selectively for predictable events. Taken together, these effects suggest that cortical knowledge representations in the default mode network and elsewhere might exert a top-down influence on memory processing in the hippocampus (and sensory cortex), potentially by stabilizing encoding activity.

Although our experimental paradigm differs considerably from the majority of studies in this space, our results are consistent with multiple past findings. We found that the predictable action sequence was represented across major nodes of the default mode network (i.e., mPFC and PMC), which have previously been shown to represent other forms of structured knowledge, such as schemas and event models (Baldassano et al., 2018; Masís-Obando et al., 2022; Reagh & Ranganath, 2023), concept or category structure (Audrain & McAndrews, 2022; Mack et al., 2020; Sommer et al., 2022), and social information (Raykov et al., 2021). Our findings also fit with models of the role these regions play in prior knowledge-mediated learning, which argue that representations in mPFC and posterior cortical regions can provide a kind of “template” (or scaffold) during novel learning, biasing attention toward context-relevant information and enabling the integration of new information with existing memories (Gilboa & Marlatte, 2017; Preston & Eichenbaum, 2013). Beyond these regions, premotor cortex also represented the predictable action scaffold, consistent with previous evidence for its role in motor sequence learning (Kornysheva & Diedrichsen, 2014; Yokoi & Diedrichsen, 2019).

In contrast, encoding activity in the hippocampus and LOC did not reflect the predictable action sequence *per se*. Upon reflection, these null effects are not particularly surprising. Although the hippocampus can represent schematic information (Mack et al., 2016; Masís-Obando et al., 2022) as well as learned sequential structure (Hsieh et al., 2014; C. Liu et al., 2021), it is implicated more strongly in the processing of novel episodic content (Sommer, 2017; van Kesteren et al., 2012). Additionally,

although Guo and Yang (2023) identified schema representations in LOC, the schema in their task (a grid of object-location associations) explicitly encompassed information about visual objects. In our task, conversely, the objects that participants encoded were unrelated to their prior action sequence knowledge. However, despite the fact that neither the hippocampus nor LOC showed predictable template activation, predictable encoding events were nevertheless associated with context-specific activity patterns in both regions — with LOC also representing variable encoding events in a selective manner. Several features of our task aside from learned motor sequence knowledge may have driven cross-event similarity in these ROIs. Namely, encoding events from the same context shared the same store identity, similar perceptual features (i.e., a static background image of the store’s interior), and the same category of objects (i.e., food or animals).

One puzzling finding — also related to the representation of motor sequence knowledge — is that the observed association between template activation and subsequent order memory performance was not selective to predictable encoding events. Specifically, in mPFC, PMC, and premotor cortex, activation of the variable template during variable encoding events was also linked to better order memory. This result implies that the variable sequence template contained meaningful contextual information, despite the fact that it was not associated with a consistent sequence of actions. In our experiment, both the predictable and variable contexts conferred some degree of prior knowledge: participants were familiar with the identity of both stores, their unique perceptual attributes (i.e., the background image of each store interior), and the general structure of both event types (i.e., that they would involve six sequential motor actions). The crucial difference between the two conditions was that the predictable store *also* afforded knowledge about a specific embedded action sequence. Our findings suggest that irrespective of whether a context allows for precise temporal predictions, the neural representation of context-specific knowledge during encoding (which is putatively captured by the predictable and variable templates) is beneficial for temporal order memory. Critically, the execution of a predictable action sequence *further* enhances memory performance, potentially via the stabilization of brain activity during encoding. Although previous studies have utilized many different kinds of prior knowledge to understand their influence on

novel learning, few if any have directly contrasted the effects of more specific versus general knowledge structures. Exploration of these differential effects is thus a worthwhile avenue for future research.

Another notable result from this study is that the stability of neural activity across time — again in both default mode and sensorimotor regions — was associated with temporal order memory. Theories of temporal context posit that the brain maintains a mental context representation that gradually drifts across time; items encoded nearby in time are associated with similar contexts, which increases the probability that those items will be remembered in serial order during recall (Howard & Kahana, 2002; Polyn et al., 2009). Relatedly, event segmentation theories argue that items perceived to belong to the same event are bound together in memory due to their shared contextual features (Clewett & Davachi, 2017; DuBrow et al., 2017). In line with these ideas, here we showed that within an episodic encoding event, greater similarity between sequential representations — potentially indicating stronger temporal context stability — was beneficial for subsequent order memory. This effect was present in both of our conditions, suggesting that it also functions as a general mechanism of episodic sequence encoding. Several previous studies have reported similar observations, particularly within the hippocampus. The stability of item representations in the hippocampus has been shown to correlate with better order memory and compressed estimates of temporal distance (DuBrow & Davachi, 2014; Ezzyat & Davachi, 2014). Hippocampal stability across time also increases as participants become more familiar with the temporal structure of an event (Kalm et al., 2013; Paz et al., 2010).

In contrast to our results, some studies have found that successful order memory is driven by greater neural *instability* (i.e., pattern dissimilarity) during sequence encoding in both the hippocampus and mPFC (Jenkins & Ranganath, 2016), and that learning increases temporal differentiation within events in the hippocampus (Bellmund et al., 2022). This work argues that more distinctive neural representations might prevent proximal items from blurring together in memory via a temporal pattern separation mechanism (Kesner, 2013; Yassa & Stark, 2011). As others have pointed out (Bellmund et al., 2022; DuBrow & Davachi, 2017), whether pattern stability or instability benefits temporal memory may depend on many factors, including how exactly memory is tested or what strategies people use during

encoding. In our experiment, we tested order memory by showing participants all six items from a recent event and asking them to select items in the original order, whereas other studies (e.g., Jenkins & Ranganath, 2016) have presented participants with two items at a time (from a longer list of items) and asked them to judge which was more recent. Different test formats might incentivize different strategies for encoding sequential information, which may manifest as different patterns of brain activity. Different brain regions within the memory system might also be biased more strongly toward temporal stability versus differentiation. For example, one recent study (Bein & Davachi, 2024) found that activity patterns in hippocampal subfield CA3 were similar for items within the same sequential event, whereas in the dentate gyrus, adjacent sequence items were represented more dissimilarly than distant items. Ultimately, future work will be necessary to clarify the features that shape how context stability relates to episodic event memory, as well as how context-sensitive representations differ across brain regions.

We hypothesized that one mechanism whereby specific motor sequence knowledge (i.e., a predictable action scaffold) might facilitate temporal memory for unrelated object sequences is by boosting contextual stability during encoding events. Several findings in our study support this hypothesis. First, in PMC, within-event stability was significantly greater in predictable versus variable events. We then found that stronger multivariate representations of the predictable action scaffold in PMC and premotor cortex were linked to stability in the hippocampus and LOC, specifically during predictable encoding events. This finding in particular is reminiscent of previous evidence that univariate activity in mPFC — which the authors interpreted to reflect schematic event knowledge — was related to greater persistence (i.e., stability) of LOC activity patterns within an event (Ezzyat & Davachi, 2021). As mentioned above, these effects jointly suggest that cortical knowledge representations can modulate brain activity during encoding, particularly in regions that are more directly involved in the encoding of novel stimuli.

Our data indicate that a diverse and distributed network of brain regions represent structured knowledge and/or temporal context during our task, including areas implicated in both domain-general and stimulus-specific processing. The task features that drove these representations likely vary across

ROIs. Context-selective representations in LOC might have emerged from the shared category or perceptual features of objects encoded in each store (Grill-Spector et al., 2001; Sayres & Grill-Spector, 2008), whereas those in premotor cortex stemmed from the motor actions participants planned and executed (Crowe et al., 2014; Kornysheva & Diedrichsen, 2014). Interestingly, across all of our analyses, activity in PMC was the most consistently modulated by the predictable action scaffold, whereas effects in mPFC were similar across the predictable and variable contexts. Based on existing literature, it is unclear how exactly mPFC and PMC differ in their informational content and their roles in prior knowledge-mediated memory. One speculative possibility is that while both regions can represent schematic knowledge, PMC is tuned to more specific or inflexible kinds of structure (e.g., a fixed motor action sequence), whereas representations in mPFC are comparatively more abstract. This idea is loosely compatible with other recent findings: for example, Reagh and Ranganath (2023) found that PMC activity differentiated between similar spatial contexts associated with the same schema while mPFC did not, and Masís-Obando et al. (2022) found that successful memory for narrative stimuli was linked to abstract schema representations in mPFC during encoding but to event-specific representations in PMC during retrieval. Another possible distinction between these regions arises from the “PMAT” framework (Ritchey et al., 2015), which argues for the existence of two cortical networks that represent different types of memory content: a posterior-medial (PM) portion (which includes PMC) that is selective for contextual information, and an anterior-temporal (AT) portion that represents items or concepts. Although mPFC is sometimes lumped within the PM network, it may also function as an integrative hub between AT and PM regions (Ritchey et al., 2015), perhaps imbuing it with greater representational flexibility. Future work that systematically varies the specificity of prior knowledge and/or event structure will be necessary to test these ideas.

One aspect of our methodology worth highlighting is the approach we used to identify neural representations of the predictable action scaffold. Despite the fact that the scaffold is inherently sequential, we defined the predictable sequence template as a single compressed pattern reflecting participants’ average brain activity as they executed the predictable action sequence. This averaging

approach mirrors how event representations are typically extracted in paradigms that use movies or spoken narratives as stimuli (Baldassano et al., 2018; Chen et al., 2017; Zadbood et al., 2017). In a supplemental analysis, we found that this compressed predictable template remained active throughout most of each encoding event, indirectly suggesting that the structure it captured remained relevant across time. However, it is possible that the predictable action scaffold is also represented in a more temporally dynamic manner, e.g., by a set of patterns that correspond to each individual motor response or sequence position. Given the modest temporal resolution of fMRI, resolving fine-grained sequential representations is difficult. Our attempt to extract them using one popular approach, Hidden Markov Models (Baldassano et al., 2017), has also been largely unsuccessful. Nevertheless, other methodological innovations, such as probabilistic pattern classifiers (Wittkuhn & Schuck, 2021), may still prove effective in characterizing how the brain moves through sequential states in the present task. An intriguing question is whether holistic memory for the order of items within an event is better predicted by contextual representations that are sustained throughout an event or by those that evolve (in a structured way) across time.

As highlighted in our previous behavioral work (Gasser & Davachi, 2023; Chapter 1), it is unclear whether the temporal memory enhancement we observed stems specifically from motor behavior. Instead, it could be that any kind of salient sequential structure — irrespective of whether it involves actions — benefits order memory. Even in our task, the action sequence that participants learned was not purely motor, given that each motor response was linked to a different spatial location. Past research provides evidence that motor actions in particular facilitate the encoding of individual stimuli (Engelkamp et al., 1994; Kinder & Buss, 2021; Madan & Singhal, 2012; Noroozian et al., 2022; Yebra et al., 2019), and that motor sequence learning can influence performance on an asynchronous list learning task (Moshay & Robertson, 2016; Mutanen et al., 2020). However, temporal memory benefits have also been seen when participants encode item sequences alongside musical stimuli with predictable structure (McElhinney & Annett, 1996; Ren et al., 2024, but also see Rainey & Larsen, 2002). Relatedly, the “Method of Loci” mnemonic strategy — which can boost serial recall performance (McCabe, 2015) — involves deliberately associating a novel list of stimuli with locations on a highly familiar spatial route (e.g., a walk through

your childhood home). These latter findings suggest that sequence knowledge across multiple domains can aid concurrent learning, while also highlighting the need for additional investigation into how and when different kinds of prior knowledge structures facilitate different forms of episodic memory.

To conclude, in this study we explored multiple neural mechanisms that facilitate the construction of temporally structured event memories, as well as how such mechanisms are influenced by prior sequential knowledge. As part of this exploration, we found that remembering an event sequence was facilitated both by knowledge representations and by temporal stability during encoding, and that learned action sequence knowledge modulated these effects — particularly within the posterior medial cortex (PMC). This work connects to research from multiple subdomains of cognitive neuroscience, including the relationship between prior knowledge and novel encoding, interactions between motor and episodic memory systems, and the process by which we remember the temporal structure of experience. Ultimately, the way that events unfold across time is a fundamental dimension of episodic memory (Eichenbaum, 2017; Tulving, 1985), essential to our ability to make sense of the past and predict the future. Investigating how temporal memory interacts with other aspects of our experience — including active motor behavior and prior knowledge — is therefore critical to our understanding of how we learn amidst the chaos and complexity of everyday life.

**Chapter 3: Novelty in everyday life promotes long-term memory for  
autobiographical events**

Camille Gasser, Victoria Schelkun, Kathryn Lockwood & Lila Davachi

### 3.1 Introduction

Novelty is among the most salient dimensions of human experience. Something new or unexpected — like your first day of high school, or the morning you saw a celebrity at your local coffee shop — will often capture your attention and stay vivid in your memory for years to come. A substantial body of previous work has investigated the effects of novelty on episodic memory (Kafkas & Montaldi, 2018; Lorents et al., 2023; Quent et al., 2021; Ranganath & Rainer, 2003; Schomaker, 2019). Whereas many findings validate the intuition that a novel experience is easier to remember (Åberg & Nilsson, 2003; Habib et al., 2003; Kormi-Nouri et al., 2005; Tulving & Kroll, 1995), other work provides evidence that familiarity too can bolster learning (Poppenk et al., 2010; Reder et al., 2013, 2016). This apparent contradiction is likely accounted for by many factors, including variation in the type of novelty experienced and the type of memory tested (e.g., memory for the gist of an event vs. specific details).

Importantly, the respective benefits of novelty and familiarity on memory are thought to be underpinned by distinct neurocognitive mechanisms. Detection of a novel stimulus evokes a cascade of neuromodulatory responses that can bolster memory by upregulating hippocampal encoding processes (Duszkiewicz et al., 2019; Kafkas & Montaldi, 2018; Schomaker & Meeter, 2015). In contrast, memory for a familiar stimulus can benefit from congruence with prior knowledge structures (e.g., schemas), which act as “scaffolds” for incoming episodic information and allow new memory traces to become rapidly integrated within pre-existing associative networks (Gilboa & Marlatte, 2017; McClelland, 2013; Preston & Eichenbaum, 2013; van Kesteren et al., 2012). From this background, it is clear that both the novelty and the familiarity of an experience can aid learning. However, with some exceptions, the vast majority of studies in this area have examined these variables in controlled laboratory experiments, limiting the generalizability of their conclusions. Here, we conducted an intensive longitudinal “daily diary” study (Bolger et al., 2003; Larson & Csikszentmihalyi, 2014) that tracked participants across multiple weeks while recording rich information about their day-to-day experiences. Using this dataset, we build upon previous literature by exploring how novelty (or the lack thereof) in everyday life impacts

multiple forms of autobiographical memory.

Understanding the impact of novelty on memory has taken on new significance in light of the recent COVID-19 pandemic. Across the world, lockdown protocols and stay-at-home orders narrowed the scope of our worlds, sharply reducing the novelty and variety of our regular activities. In anecdotal reports from this period, people often lamented that their memory felt blunted or distorted, with days blurring together and very few moments standing out as vivid or memorable (Castellà et al., 2024; Chaumon et al., 2022). This phenomenon can be contextualized within multiple theories describing how we remember episodic events. Theories of event segmentation posit that stability and change in one's context (e.g., in one's surroundings or mental state) help to segment continuous experience into discrete and meaningful "episodes" (Clewett et al., 2019; Clewett & Davachi, 2017; Shin & DuBrow, 2021). Previous work has found that such segmentation is beneficial for long-term memory (Kurby & Zacks, 2008; Pettijohn et al., 2016; Sargent et al., 2013). As such, it follows that a lack of change in one's environment might impair memory function by dampening this organizational process. Although event segmentation is typically described as unfolding across the span of minutes or hours (but see Yousif et al., 2024), similar ideas are discussed at a larger scale by Brown's (2016) Transition Theory of autobiographical memory. This theory argues that extended periods of stability in one's life (e.g., living in the same apartment with the same job for a year) create "lifetime periods" that are separated by major transitions (e.g., moving to a new city). Whereas memories from the middle of a lifetime period will be relatively scarce, transitions function as "temporal landmarks" that facilitate access to memories near that time (N. R. Brown, 2021; Shum, 1998). Taken together, these theories suggest that a lack of novelty or change in one's day-to-day life might reduce the quality or availability of memories from that period. When examining the effects of novelty on memory, therefore, it is critical to consider not only the novelty of an event itself, but also the novelty embedded within one's broader experience.

A separate body of work provides further insight into how novelty can exert diffuse effects on memory. Across multiple paradigms, researchers have found that exposing animals to a new spatial environment shortly before *or* shortly after a brief learning task improves long-term memory (Li et al.,

2003; Moncada & Viola, 2007; Wang et al., 2010). Similar effects have since been observed in humans (Fenker et al., 2008; Schomaker et al., 2014), including multiple studies held in real-world classroom settings showing that students who engaged in a novel activity shortly before or after an unrelated memory task performed better on that task than those who did not (Ballarini et al., 2013; Ramirez Butavand et al., 2020). The temporal window of this “novelty penumbra” effect is still the subject of some debate: whereas some findings suggest that the benefits of novelty extend only to events taking place within 30 to 60 minutes (e.g., Ballarini et al., 2013; Li et al., 2003), others indicate that a novel experience up to 11 hours after learning can still benefit memory (Ramirez Butavand et al., 2020). But on the whole, this line of research provides further evidence that engaging in new activities can bolster the strength or longevity of episodic memory, even for experiences that lack novelty on their own.

The last decade of research has seen an increase in the number of studies that seek to characterize learning and memory in more naturalistic settings. Much of this work asks participants to document moments from their day-to-day lives by taking pictures or videos (Bainbridge & Baker, 2022; Chow & Rissman, 2017; Doherty et al., 2012; Martin et al., 2022; Nielson et al., 2015) or generating diary entries (Burt et al., 2003; Pearson et al., 2023; Rasmussen et al., 2014), allowing researchers to probe memory for personal autobiographical experiences. Several recent studies are of particular relevance to the present work. Rouhani et al. (2023) and Folville et al. (2023) both examined participants' memory for experiences during and adjacent to the COVID-19 pandemic. Folville et al. (2023) focused on the relationships between subjective well-being, experiential diversity (a construct closely related to novelty) and episodic memory, finding that individuals who engaged in more diverse activities maintained better well-being during pandemic lockdowns; however, experiential diversity was not directly related to more vivid memories for lockdown experiences. Somewhat contrary to these effects, Rouhani et al. (2023) found that participants were more likely to spontaneously recall events from March 2020 than from other months that year — suggesting that a highly novel and historic event, like the onset of a major pandemic, can heighten memory for nearby experiences (before extended lockdown periods presumably dampen it). In another study, Bainbridge and Baker (2022) queried autobiographical memory

using 1-s videos that participants had previously recorded using a smartphone app, finding that memory was better when videos contained unfamiliar people or were filmed in unfamiliar locations.

In the current work, rather than using global events to infer moments of novelty or relying on retroactive judgments of past experience, we asked participants to complete daily diary entries for two weeks, during which they: (1) recorded information about the novelty of their everyday experiences (e.g., how typical each day felt, whether they visited a new location); and (2) provided written descriptions of multiple specific events they had engaged in each day. Critically, we then explored the relationship between memory for these autobiographical experiences and the novelty experienced throughout the diary period. We focused in particular on two forms of memory: memory for individual episodic events, and memory for the temporal relationship (i.e., order and distance) between different events.

### **3.2 Methods**

#### **3.2.1 Participants**

A total of 51 participants were enrolled in the current study from the Columbia University community. Data collection took place in two recruitment waves: Group A ( $N = 19$ , between 11/23/21 and 12/27/21) and Group B ( $N = 32$ , between 3/22/22 and 5/3/22). All participants were between the ages of 18 and 35, lived in the United States, and reported having normal or corrected-to-normal vision. They also indicated that they were not currently diagnosed with any neurological, neurodevelopmental, psychotic, dissociative, substance-related, neurocognitive, or personality disorder. From this sample,  $N = 10$  were withdrawn from the study because they missed at least two entries during the two-week daily diary period. As such, we were left with a final sample of  $N = 41$  participants ( $M_{age} = 25.6$ ;  $SD_{age} = 4.2$ ; 82.9% female). Participants were paid at a rate of \$12/hour for all time spent completing study tasks. All study procedures were approved by the Institutional Review Board at Columbia University.

#### **3.2.2 Study design**

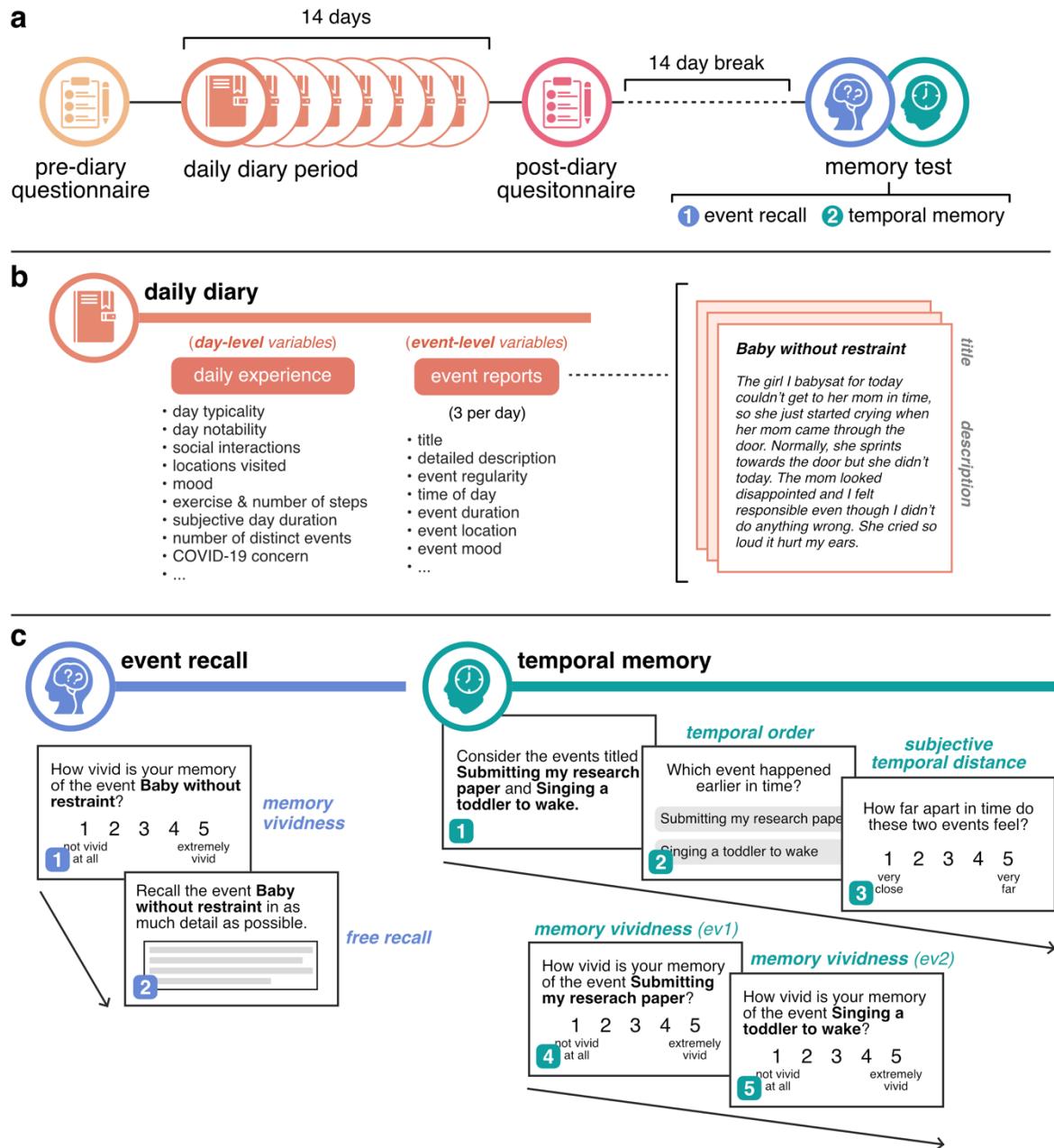
This study adopted an intensive longitudinal “daily diary” design, in which participants completed a sequence of study tasks across a 4-week period. All portions of the study were administered

remotely via Qualtrics (<https://www.qualtrics.com>), and could be completed on any web-enabled device (e.g., computer, tablet, phone). The study began with the *pre-diary questionnaire*, which included demographic questions, assessments of affective and emotional states, questions about attitudes toward COVID-19, and questions about participants' daily experiences and routines. The day after completion of the pre-diary questionnaire, participants entered the 2-week *daily diary period*. At the end of each day within this period, participants were prompted to complete a daily diary entry that collected information about their experiences, behaviors, and feelings on that day (e.g., how many places they went, who they interacted with). They were also asked to report information about three distinct events or activities they had engaged in that day. After the last day of the diary period, participants completed the *post-diary questionnaire*, which was nearly identical to the pre-diary questionnaire in its content (see below for details). Finally, after a 2-week delay, we administered a set of *memory tests* that probed participants' memory for the specific events they had reported as part of the diary period. See Figure 3.1a for a schematic of the entire study structure.

### ***Pre-diary questionnaire***

The pre-diary questionnaire collected a rich array of background information about each participant's demographics, current and overall emotional state, typical daily routine, and feelings toward the COVID-19 pandemic. Demographic questions asked about participants' age, gender, race, education level, and socioeconomic status. To assess emotional well-being, we administered a variety of validated psychological inventories: namely, the State-Trait Anxiety Inventory (STAI-T; Spielberger et al., 1983), the Positive and Negative Affect Scale (PANAS; Watson et al., 1988), the Intolerance of Uncertainty Scale (IUS; Freeston et al., 1994), and the Depression, Anxiety, and Stress Scale (DASS-21; Henry & Crawford, 2005). Questions related to COVID-19 assessed participants' concern with themselves and/or close others contracting COVID, predictions about the future state of the pandemic, stress in response to the effects of the pandemic, and consumption of COVID-related news. Participants also reported information about their typical daily experience, such as descriptions of their everyday routine, the number of places they generally visited each day, the number of social interactions they engaged in each

day, and so on. Although we do not consider data from the pre-diary questionnaire as part of the current manuscript, see Appendix A, Table 1 for an overview of information collected during this task.



**Figure 3.1. Study design.** (a) Illustration of each phase of the study. The two-week daily diary period was flanked by the administration of two questionnaires (“pre-diary” and “post-diary”). Two weeks after the post-diary questionnaire, participants completed a memory test for events recorded during the diary period. (b) Overview of variables collected during each daily diary (see *Methods* for details). Importantly, on each diary day, participants reported information about three distinct events from that day, providing a self-generated title and written description for each. (c) Illustration of the two memory test phases: event recall (left) and temporal memory (right). Note that both memory tests collected subjective memory vividness ratings for each individual event included in the test.

### **Daily diary period**

The day after the pre-diary questionnaire was completed, participants entered the two-week daily diary period (Figure 3.1b). At 5pm on each day, they received an email with a link to that day's diary survey, which they were instructed to complete at any point between receipt of the link and when they went to bed (*average start time* = 8:39pm). The median time taken to complete the diary was 18.8 minutes, although this time varied considerably across participants ( $SD = 38.7$  minutes). Participants were informed during recruitment that missing more than two daily diaries would end their study participation. The majority of our sample ( $N = 27$  of 41) completed all 14 diaries, with a smaller subset missing one ( $N = 13$ ) or two ( $N = 3$ ) days. Due to a technical error, one participant completed one extra daily diary, and thus had diary data from 15 days.

Each daily diary began with a set of questions about participants' thoughts, feelings, and actions during that day. For example, participants were asked to report information about their current mood, the places they had visited outside of the home, their exercise levels, and their social interactions that day. Then, they were prompted to report information about three specific "events" that had occurred on that day. In the task instructions, we defined an event broadly as "a thing you have done (e.g., listening to music, working on a project, talking on the phone, etc.), either on your own or with others." Participants provided a written text description of each individual event, and were encouraged to be as detailed as possible. Next, they were prompted to assign a unique title to that event that was relatively brief, but distinctive enough to refer to the specific experience in question. The remaining set of event-specific questions collected information about the subjective and objective duration of the event, when it took place, other people who were present during the event, how often the participant had experienced that kind of event, how memorable and meaningful it was, what emotions it evoked, and what took place immediately before and after. Participants described three specific events per daily diary. Finally, after reporting these events, they answered a short set of questions about their consumption of news that day (e.g., what stories they heard, whether they discussed the news with others). Given the large amount of information collected as part of these daily diaries, this paper does not include details about every

collected variable — but see Appendix C, Tables 2 and 3 for a complete list of all diary questions, including specific text prompts and response formats.

The primary motivation for the present study was to understand how the experience of novelty in daily life shapes episodic memory for real-world, autobiographical events. Critically, these daily diaries provided us with: (1) detailed information about the events that participants experienced (including detailed text descriptions and self-generated titles/cues); and (2) an array of variables that reflected the novelty of the event itself, as well as other novel happenings throughout that day. For details about which specific diary items were used to quantify novelty, see the section below titled *Measuring novelty during the diary period*.

#### ***Post-diary questionnaire***

The post-diary questionnaire was administered the day after the end of the diary period, and was intended to capture any changes in participants' mental states and/or typical routines while they were enrolled in the study. This questionnaire was identical to the pre-diary version, except that it did not contain questions about demographic background. See Appendix C, Table 1 for a summary of questionnaire items.

#### ***Memory tests***

The memory test portion took place roughly two weeks after the conclusion of the daily diary period ( $M_{delay} = 16.1$  days;  $SD_{delay} = 1.2$  days;  $range = 14\text{-}19$  days), and was organized into two sections: (1) event recall, and (2) temporal memory. During the event recall portion (Figure 3.1c, *left*), participants had to recall detailed information about the events they had recorded several weeks prior. Events were cued using the titles generated by the participants themselves during the diary period. For each event, participants first provided a *memory vividness* rating, which served as a measure of the subjective strength with which they could recall the specified event. They then attempted to *freely recall* the event in as much detail as possible. Following these two memory reports, participants were cued to recall specific pieces of information about the event in question, including the location and time of day in which it took place, any others who were present during the event (e.g., friends, family members), how long the event lasted, and

what they were doing immediately before and after the event. They also provided confidence ratings for each of these items. See Appendix C, Table 4 for details about the language and response format for each question in this test.

Next, participants completed the temporal memory test (Figure 3.1c, *right*). On each trial, they were first cued with the titles of two events from the diary period. Next, they indicated which of the two events they remembered having taken place earlier in time (*temporal order memory*), as well as how confident they were in that judgment. We then collected measures of the *subjective temporal distance* between the two events (i.e., how far apart in time participants felt the events to be), as well as their *objective temporal distance* (i.e., how many actual days separated the two events). Next, participants indicated how similar they felt the two events were to each other. Finally, they provided *memory vividness* ratings for each individual event in the pair — analogous to those provided for each event included in the event recall test. Again, see Appendix C, Table 4 for details about each test question.

As described above, each participant reported up to 42 distinct events throughout the full diary period ( $M = 40.2$  events;  $SD = 2.5$  events). One third of these events ( $M = 13.8$  events;  $SD = 0.6$  events), one per diary day, were randomly selected to include in the event recall test, while the remaining two-thirds were used to create event pairs for the temporal memory test ( $M = 13.2$  pairs;  $SD = 1.1$  pairs). We created event pairs by pseudo-randomly selecting events that were reported either two or three days apart during the diary period. Event pairs with two-day and three-day lags were sampled in roughly equal proportions within each participant, with 51.4% of all pairs separated by two days and 45.5% separated by three. However, due to a coding error in the memory test generation process, a small number of event pairs were instead separated by one ( $N = 5$  pairs, or 0.9%) or four ( $N = 11$  pairs, or 2.3%) days. Importantly, both phases of the memory test (event recall and temporal memory) asked participants to rate the vividness of their memory for each event described. Thus, although each event was included in either the event recall or the temporal memory test, we obtained memory vividness ratings for all events reported during the study.

### **3.2.3 Measuring novelty during the diary period**

In order to examine how novelty in everyday life impacts autobiographical memory, we identified a set of variables collected as part of the daily diaries that were indicative of new or atypical experiences. We focused on five variables: (1) *event regularity*; (2) *day typicality*; (3) *novel locations visited*; (4) *novel social interactions*; and (5) *day notability*. The first variable — *event regularity* — reflected the novelty of the reported events themselves. Specifically, during daily diaries, participants assigned each event to one of three categories: “routine” (something that they do almost every day), “periodic” (something that they do occasionally), or “new” (something they had never done before) (Appendix C, Table 3: Q14). The remaining four variables assessed novelty at the level of diary days. Note that these variables do not allow us to draw precise conclusions about when exactly during the day a new experience took place, but nevertheless give us broad insight into what kinds and the amount of novelty participants were exposed to each day. First, ratings of *day typicality* were collected by asking participants to make a subjective report of how similar the current day was to the “typical day” they had described as part of the pre-diary questionnaire (Appendix C, Table 2: Q4). *Novel locations visited* was a binary variable that denoted whether the participant had reported going to any kind of novel location that day, either as part of one of the events they had described (Appendix C, Table 3: Q10) or while providing information about all locations they had visited that day (Appendix C, Table 2: Q9). Similarly, *novel social interactions* was a binary variable that indicated whether the participant had described interacting with a new person as part of an event and/or at any point on that day (Appendix C, Table 2: Q15/16 & Table 3: Q6/8). Finally, *day notability* was a binary measure obtained by asking whether anything “notable, out of the ordinary, very important, and/or unexpected” had happened on that day (Appendix C, Table 2: Q20). These measures of novelty were then connected to participants’ memory for events reported throughout the diary period.

### **3.2.4 Measuring episodic memory for individual events**

Throughout this chapter, we examine two measures that assess autobiographical memory for individual events. The first — *memory vividness* — reflects the subjective strength of a participant’s

memory for the cued experience, and was reported on a 1-5 scale (with 5 indicating extremely vivid memory). The second — the number of episodic details recalled — was quantified using a modified version of the Autobiographical Interview scoring approach (Levine et al., 2002). In this method, spoken or written descriptions of past experiences are decomposed into distinct informational units, or “details.” Each detail is assigned to one of several categories, a subset of which encompass details that are episodic in nature (i.e., that refer to an individual’s perceptions, actions, and/or thoughts during the specific experience in question). Episodic details (also referred to as “internal” details) are contrasted with non-episodic (or “external”) details, which may encompass semantic information, meta-cognitive reflections on the experience, or information about events beyond the one in question. Specifically, each detail was assigned to one of five episodic categories: (1) “event” (describing what happened, who was present, etc.); (2) “place” (describing the event location), (3) “time” (describing when the event took place or how long it lasted), (4) “thought/emotion” (describing the feelings or thoughts of the self or others) and (5) “perceptual” (describing sensory experiences); or to one of three non-episodic categories: (1) “semantic” (describing general or extended states of being), (2) “reflection” (editorializing, meta-cognitive statements about the past or future), and (3) “extraneous” (describing aspects of an event other than the one in question). We quantified these details for all event descriptions provided *during the diary period*, as well as all event recalls *during the memory test*. For all analyses in this paper, we focused just on the total number of episodic details (across all five categories) per event description/recall, which we interpret to reflect the objective quantity of information recalled about each experience.

In order to segment text into discrete episodic (and non-episodic) details, we leveraged recent advances in large language models (LLMs) to automate the memory processing procedure (Michelmann et al., 2023). Specifically, we prompted GPT-4 to perform the task of parsing each memory into meaningful chunks and then categorizing each chunk to the appropriate episodic or non-episodic detail category. For each memory description, we initiated a call to the OpenAI Chat Completions API (OpenAI et al., 2024) with the model version “gpt-4-0613” (<https://platform.openai.com/docs/models/>) and provided a naturalistic description of the task instructions (see Appendix C for the exact prompt) followed

by the memory description. To increase the reproducibility, the *temperature* parameter was set to 0.1. We set the *max\_tokens* parameter to 1000 to balance between allowing sufficient resources for the model to complete the task and avoiding excessively long responses.

For comparison and verification, independent (human) raters ( $N = 3$ ) segmented and categorized a subset of the event descriptions and recalls ( $N = 657$ , or ~30% of the full sample). These raters were provided analogous instructions to the prompt used with GPT-4 for the segmentation and categorization procedure. We found a strong correlation of the episodic detail counts obtained by manual processing those produced by GPT-4 for the same events ( $r = 0.71, p < .001$ ), thus validating that the GPT-produced measures of detail counts are comparable to that which would be produced by manual raters.

### **3.2.5 Differences between participant groups**

As noted above, participants belonged to two waves of data collection: Group A participants were recruited between November and December of 2021, whereas Group B participants were recruited between March and May of 2022. Although each group completed nearly identical study procedures, there were a number of small methodological differences. The most noteworthy of these was that all Group B participants completed a brief virtual onboarding session with an experimenter prior to beginning any study tasks. During this session, the experimenter reviewed instructions that participants would see when describing events throughout the diary period. Participants were also presented with definitions and examples of what an event could refer to, useful questions they could ask themselves when recording event descriptions (e.g., “what did your surroundings look like?”, “what were you feeling?”), and examples of high-quality event descriptions and titles. This process was implemented in response to the observation that some participants in Group A generated titles and descriptions that were somewhat vague or generic (e.g., a title of “Went to work”). Note that both groups saw the same instructions as part of the daily diaries (see Appendix C, Table 3). Beyond this onboarding session, we implemented a handful of other minor changes between procedures used for Group A versus B to improve data quality. Specifically, for participants in Group B: (1) a minimum time limit of 2 minutes was enforced for the free recall of each event in the recall test; (2) we fixed a bug that was present in the daily

diary for Group A participants that prevented the collection of information about day-level social interactions (Appendix C, Table 2: Q15-17); (3) we made all diary responses mandatory; (4) we added response format validation to several diary and memory test items requiring numeric or time/date input (i.e., Appendix C, Table 2 [Q13, Q18], Table 3 [Q4], Table 4 [Q5]); and (4) the response option “family member(s)” was added when asking participants about who they had interacted throughout the day (Appendix C, Table 2: Q15) or as part of an event (Appendix C, Table 3: Q6).

Although the multilevel modeling approach used throughout this paper should account for some idiosyncrasies between participants, to address whether these procedural differences between groups impacted our results, we reran all of our primary analyses separately within each group (see Appendix C), and have indicated throughout the main text where effects hold or do not hold in each sub-group.

### **3.2.6 Individual event exclusions**

A small number of events ( $N = 13$  in total) provided by participants were excluded from all data analyses. Specifically, we removed any events that were missing titles or text descriptions, as well as events with titles that were not sufficiently distinctive for use as memory cues (e.g., from a participant with multiple events titled “Taking my dog on a walk”). For all analyses involving the number of episodic details recalled about each event, we also excluded another set of events ( $N = 44$ ) where participants used language that explicitly indicated that they forgot much or all of the experience (e.g., phrases like “I don’t remember this”). Note that keeping these events in our analyses did not change our main findings.

### **3.2.7 Statistical analysis**

The majority of analyses reported in this manuscript are based on multilevel regression models because they can capture effects both within and across individuals and have the capacity to handle data that is unbalanced across experimental conditions. For each analysis, we began by fitting a model with random intercepts and random slopes for each predictor, all nested within each participant. In the case of singular fits or convergence issues, we re-ran a simplified “random intercept-only” version of the model without random slopes. When significant effects were observed in a simplified model, we verified their

robustness by running a follow-up Bayesian regression model with the full random effects structure (see Appendix C for more details). All regression models were linear with the exception of those where temporal order accuracy was the output variable, in which case a logistic regression model was used instead. Throughout the paper, we also use two-tailed t-tests for analyses involving simple contrasts (e.g., differences between groups). For these tests, we report 95% confidence intervals for each test, as well as Cohen's  $d$  as a measure of effect size. (Note that we report all  $d$  values as positive, irrespective of the sign of the corresponding  $t$  statistic.)

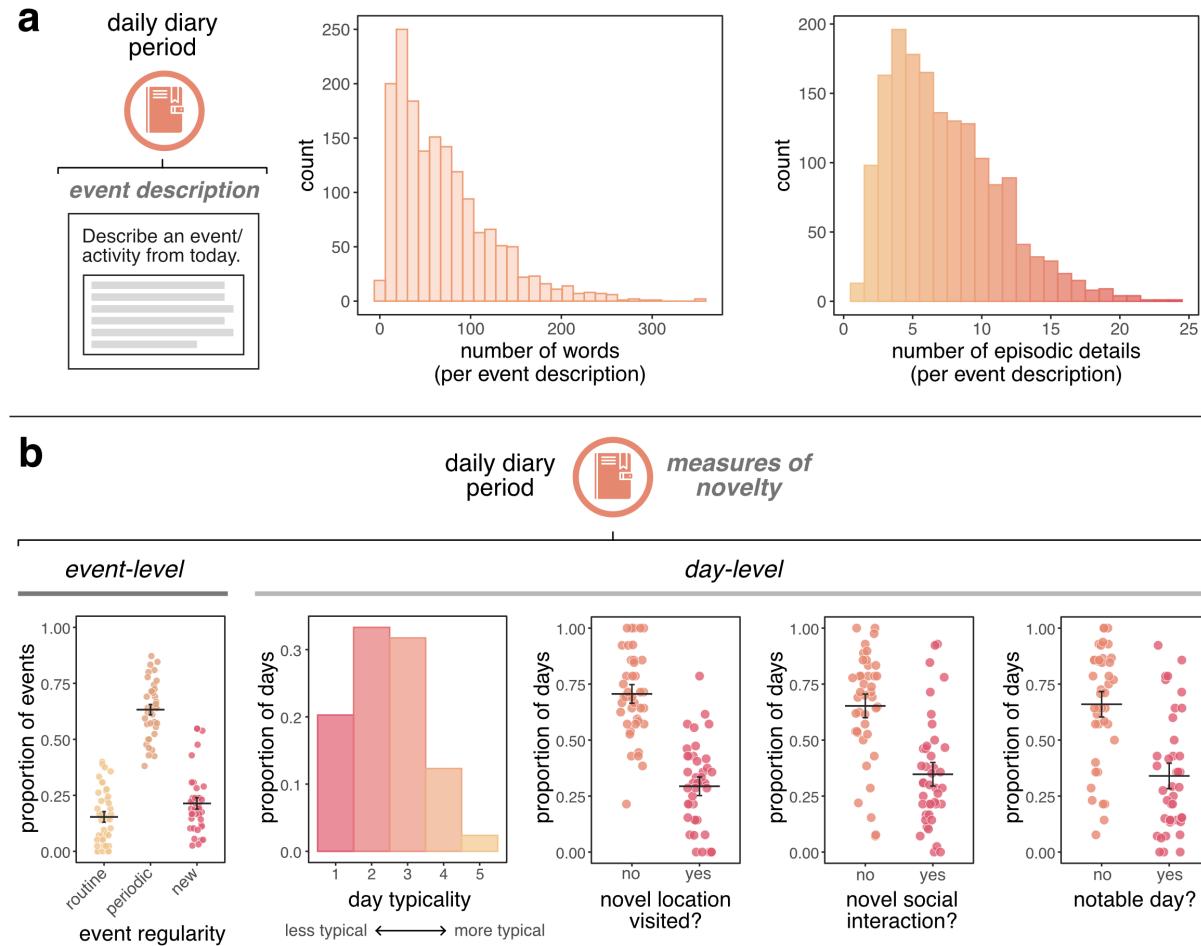
All analyses were conducted using custom analysis scripts in R (ver. 4.4.1). Functions from the *ggplot2* (ver. 3.5.1), and *sjPlot* (ver. 2.8.16) packages were used for data visualization, and the *effsize* package (ver. 0.8.1) was used for calculating effect sizes. Standard (i.e., Frequentist) regression models were computed using the *lme4* (ver. 1.1-35.5) and *lmerTest* (ver. 3.1-3) packages, with Satterthwaite's method used for computing degrees of freedom for linear models (Satterthwaite, 1946), and asymptotic Wald tests used in the case of logistic regression models. We calculated 95% confidence intervals using the *parameters* package (ver. 0.22.1) for linear models and the built-in *stats* package (ver. 4.4.1) for logistic models. When models included categorical predictors with more than two levels, the *emmeans* (ver. 1.10.3) package was used to extract estimated marginal means for each level and then run pairwise contrasts between those estimates (with  $p$ -values adjusted using Tukey's method for comparing a family of estimates; Tukey, 1949).

### 3.3 Results

#### 3.3.1 Behavior during the diary period

During the two-week daily diary period, participants reported information about three distinct events they had experienced that day. As part of event reporting, they were asked to describe in detail what happened during the event and to assign each event a unique title (Figure 3.1b). Event descriptions varied considerably in the number of words ( $\bar{M} = 69.1$ ,  $Med = 57.0$ ,  $SD = 53.0$ ) and the number of discrete episodic details ( $\bar{M} = 7.3$ ,  $Med = 7.0$ ,  $SD = 3.9$ ) they contained (Figure 3.2a). Participants in the second

wave of data collection (i.e., Group B) also tended to produce longer event descriptions than those in the first wave (Appendix C), suggesting that the procedural changes we made to increase the quality of event reporting (see *Methods*) were successful at generating more detailed responses.



**Figure 3.2. Overview of event descriptions and novelty variables.** Panel (a) includes histograms of the number of distinct words (left) and the number of episodic details (right) per event description. Episodic details were quantified using an automated scoring approach detailed in the *Methods*. Panel (b) includes summary plots of each individual novelty variable, including event regularity (i.e., whether an event was described by the participants as a routine, periodic, or new occurrence), day typicality, whether a novel location was visited that day, whether a novel social interaction occurred, and whether the day was considered notable. All error bars indicate within-subjects standard errors.

Next, we examined the set of variables we use to assess the novelty of participants' experiences throughout the diary period (Figure 3.2b). The majority of events ( $M = 63.3\%$  per subject,  $SD = 12.2$ ) were described as "periodic" occurrences (i.e., an activity that is engaged in occasionally), whereas fewer

were considered “routine” ( $\bar{M} = 15.5\%$  per subject,  $SD = 12.4$ ) or “novel” ( $\bar{M} = 21.4\%$  per subject,  $SD = 13.3$ ). With respect to our measures of “day-level” novelty (see *Methods*), we found that on average, 29.4% of days ( $SD = 18.8\%$ ) involved visiting a new location, 34.7% ( $SD = 23.8\%$ ) involved interacting with a new person, and 34.0% days ( $SD = 25.7\%$ ) were described as “notable”, in that they involved something out of the ordinary, unexpected, or important happening. Ratings of day typicality were distributed across the 1-5 scale, with an average rating of 2.4 ( $SD = 1.0$ ) across participants. Comparisons across subject groups indicated that participants in Group B reported more “new” events than those in Group A (Group A:  $\bar{M} = 14.9\%$ ; Group B:  $\bar{M} = 25.1\%$ ), but that the groups did not differ on average in the proportion of days with novel locations visited or novel social interactions, nor did participants differ in the proportion of days described as notable or in their average day typicality ratings (Appendix C). Taken together, these descriptive results demonstrate that while different types of novelty did not occur every day, participants generally reported having a considerable proportion of new experiences throughout their time in the experiment.

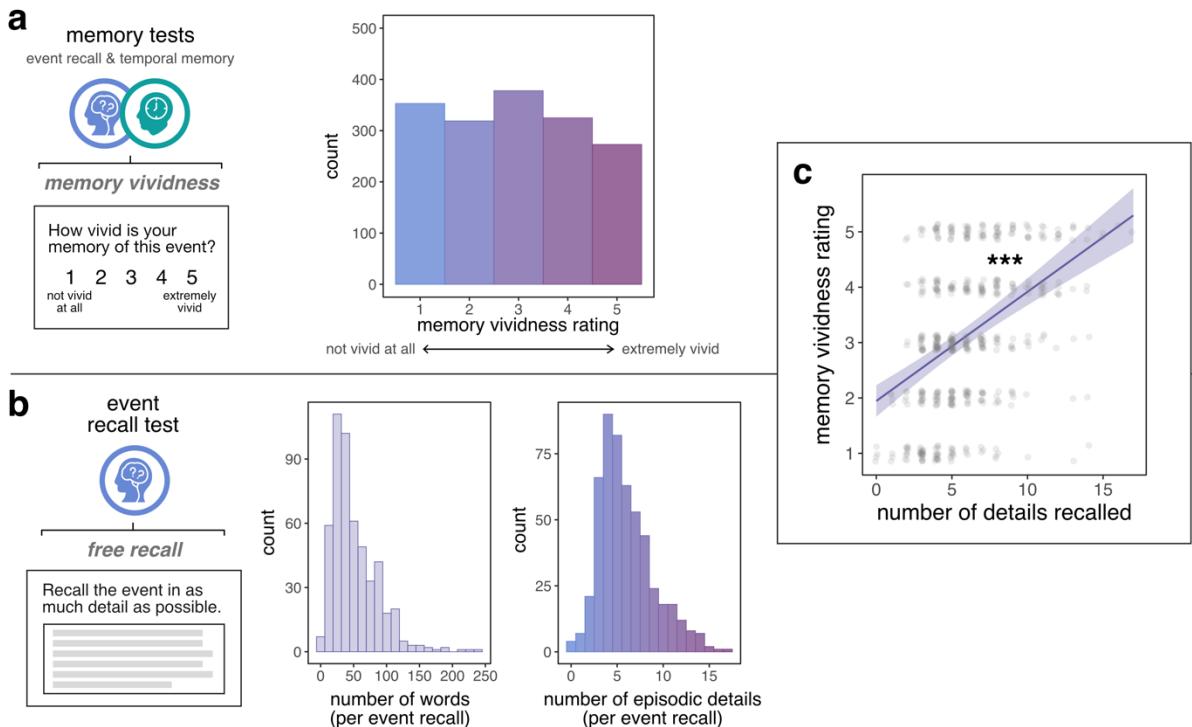
### 3.3.2 Memory performance

Before examining how novelty influenced participants’ autobiographical memory, we began by characterizing performance during our two memory tests. Both tests occurred roughly two weeks after the end of the daily diary period, and the event titles participants generated during the diary period were used as memory cues. First, we administered an *event recall* test, during which participants attempted to freely recall as much detail as they could remember about what happened during the cued event. Then, we administered a *temporal memory test*, which asked them to judge which of two events had occurred earlier in time (“temporal order accuracy”), as well as how far apart the events were perceived to be (“subjective temporal distance”). Importantly, as part of both tests, participants provided subjective “memory vividness” ratings for each individual event that was referenced (Figure 3.1c).

#### ***Event memory***

First, we examined subjective memory vividness ratings for all events. Vividness ratings followed a roughly uniform distribution (Figure 3.3a;  $\bar{M} = 2.9$ ,  $SD = 1.4$ ), indicating that participants utilized the

full 1-5 scale to reflect on the strength of their memories. Average memory vividness ratings did not differ between participants from Group A versus B (Appendix C). Because events were recorded across a multi-week period, the delay between when an event occurred and when the memory test was administered varied substantially across events, from roughly two to four weeks ( $\bar{M} = 22.6$  days,  $SD = 4.2$ ). Longer delays were associated with significantly lower vividness ratings ( $b = -0.024$ ,  $SE = 0.066$ ,  $CI = [-0.042, -0.006]$ ,  $p = 0.010$ ), an effect that was present in both participant groups (Appendix C).



**Figure 3.3. Event memory behavior.** Panel (a) shows a histogram of the subjective memory vividness ratings provided for each event as part of the memory test period, where 1 = *not vivid at all* and 5 = *extremely vivid*. Panel (b) illustrates the number of words (left) and episodic details (right) reported during the recall of each event in the event recall test. Panel (c) shows the positive association between the number of episodic details recalled about each event and memory vividness ratings. Grey dots represent individual datapoints. The purple line and shaded band indicate the fitted regression line and 95% confidence interval from a multilevel linear regression model. \*\*\* $p < .001$

Next, we considered a more objective measure of memory performance: the number of discrete episodic details that participants recalled for each event in the free recall test (Figure 3.3b). The mean number of details recalled per event was 5.9 ( $SD = 0.3$ ), with Group B recalling more details on average.

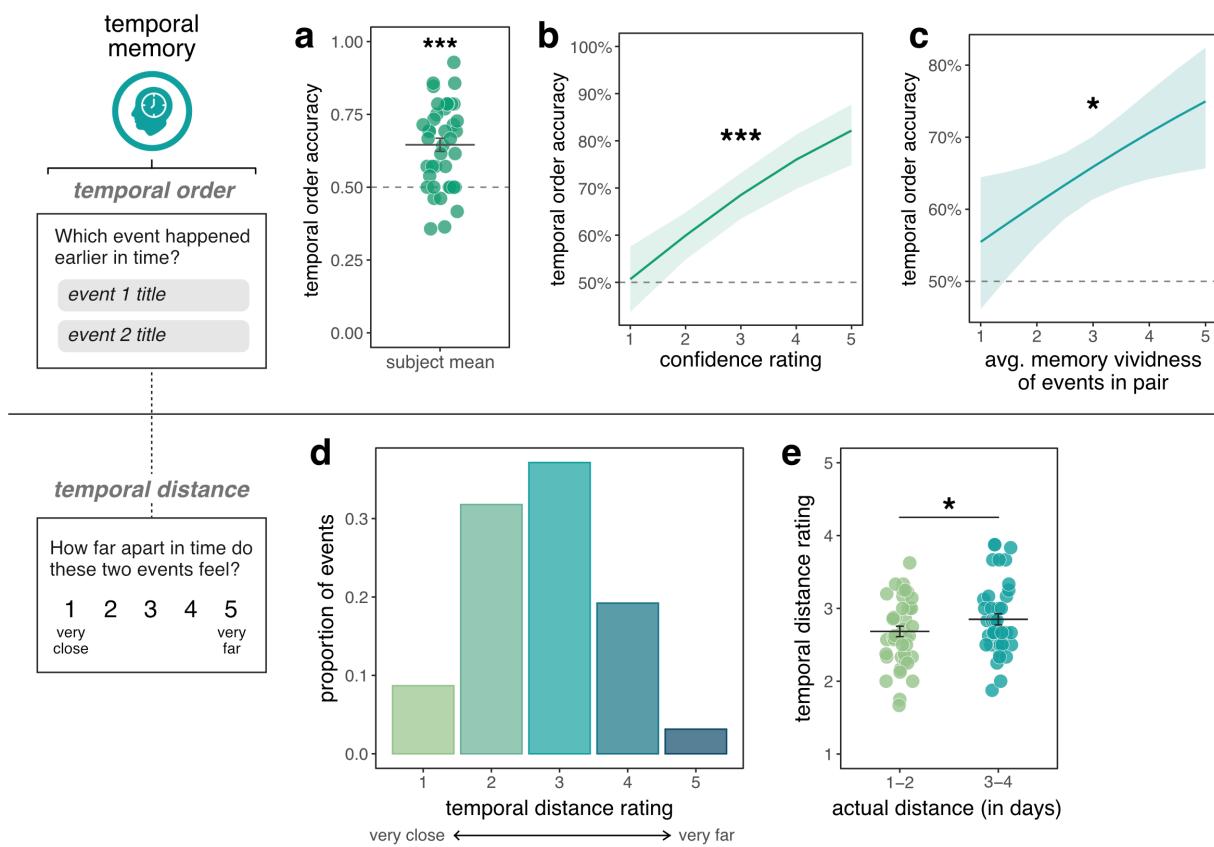
Interestingly, although participants in Group B generated more episodic details both at the time of the diary and during subsequent recall, there was a strong interaction between timepoint and group, such that the difference in the mean number of details reported by each group was lower during recall than during the diary period (see Appendix C). Of note, the number of details recalled per event was significantly associated with the number of details generated for that event *during the diary period* ( $b = 0.184$ , SE = 0.047, CI = [0.088, 0.280],  $p < 0.001$ ).

In contrast to the effect observed with memory vividness ratings, the number of episodic details recalled did not significantly differ based on the delay between the event's occurrence and the memory test ( $b = -0.030$ , SE = 0.023, CI = [-0.076, 0.016],  $p = 0.206$ ). Finally, we explored the relationship between the number of details recalled and the subjective memory vividness score for each event (Figure 3.3c). These two metrics of event memory showed a strong and significant positive association ( $b = 0.197$ , SE = 0.021, CI = [0.156, 0.238],  $p < .001$ ). This finding not only helps to validate our approach for quantifying episodic details (see *Methods*), but also to establish that subjective assessments of memory strength track with the quantifiable amount of content remembered. All associations described in this section replicated when looking separately within each participant group (Appendix C).

### ***Temporal memory***

Across all participants, temporal order memory was significantly above chance ( $t(40) = 6.465$ , CI = [0.600, 0.691],  $p < .001$ ,  $d = 1.010$ ; Figure 3.4a), and did not differ between participant groups (Appendix C). As expected, order accuracy also increased as a function of the subjective confidence ratings provided immediately after making each order judgment ( $b = 0.376$ , SE = 0.076, CI = [0.226, 0.525],  $p < 0.001$ ; Figure 3.4b). Next, we assessed the relationship between order accuracy for each event pair and memory vividness ratings for the pair's constituent events, with the logic that stronger memory for individual events should improve the likelihood that their order will be correctly remembered. Indeed, the probability of accurate order memory judgments increased based on the average vividness of the events in a given pair ( $b = 0.219$ , SE = 0.097, CI = [0.029, 0.410],  $p = 0.024$ ; Figure 3.4c). Interestingly, we did not see evidence that order memory accuracy was related to how much time had passed between

when events in the pair had occurred and the subsequent memory test ( $b = 0.015$ , SE = 0.025, CI = [-0.035, 0.065],  $p = 0.555$ ). Note that in this analysis, the day of the *first* event in each pair was used to calculate test delays, but that results were similar when using the day of the *second* event instead. All temporal order memory effects reported here replicated when looking separately in each participant group, with the exception that the relationship between order accuracy and memory vividness ratings was not significant in Group A (Appendix C).



**Figure 3.4. Temporal memory behavior.** (a) Average performance for each participant on the temporal order memory test, where participants had to indicate which of two events had occurred earlier in time. Significance stars indicate the results of a one-sample t-test comparing order accuracy to chance level (0.5). (b) Relationship between confidence ratings provided for each temporal order judgment and accuracy. (c) Relationship between the average memory vividness scores for both events in a pair and the accuracy of order judgments. Dashed lines in panels (b) and (c) indicate chance performance. Green and blue lines and shaded bands indicate the fitted regression lines and 95% confidence intervals from multilevel logistic regression models. (d) Distribution of subjective temporal distance ratings for events in each pair. (e) Average temporal distance ratings for event pairs that were separated by 1-2 days versus 2-3 days. Error bars in (a) and (e) indicate within-subjects standard errors. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

We then turned to subjective judgments of temporal distance, in which participants rate how far apart in time the events in each pair feel on a scale from 1 (“very close”) to 5 (“very far”). Across all event pairs, temporal distance ratings followed a roughly normal distribution (Figure 3.4d), with participants in Group B reporting longer distances on average (Appendix C). We then checked whether subjective distance ratings varied based on objective temporal distance. Because the majority of pairs contained events that occurred either two or three days apart — with only a handful separated instead by one or four days (see *Methods*) — we compared the mean subjective distance for event pairs that were 1-2 days apart versus 3-4 days apart (Figure 3.4e). As expected, subjective distance ratings were higher for pairs separated by more days ( $t(40) = -2.240$ ,  $CI = [-0.316, -0.016]$ ,  $p = .031$ ,  $d = 0.350$ ). When examining each group separately, this effect was only significant in Group A (Appendix C).

### 3.3.3 Effects of novelty on event memory

Our driving motivation for the present study was to understand how novelty and familiarity in everyday life affect memory for autobiographical experiences. To this end, we explored relationships between our collection of novelty variables (measured during the two-week daily diary period; see *Methods* and Figure 3.2b) and our two indices of memory for individual events: memory vividness scores, and the number of episodic details recalled. We began by asking how memory differed based on the regularity of the events themselves, i.e., whether the event was described (by the participant) as something “routine”, “periodic”, or “new.” We found that memory vividness ratings scaled strongly with the degree of reported novelty (Fig. 5a), such that new events were recalled more vividly than periodic events ( $b = 0.801$ ,  $SE = 0.099$ ,  $CI = [0.561, 1.042]$ ,  $p_{tukey} < .001$ ) and routine events ( $b = 1.622$ ,  $SE = 0.135$ ,  $CI = [1.290, 1.953]$ ,  $p_{tukey} < .001$ ), and periodic events were recalled more vividly than routine events ( $b = 0.820$ ,  $SE = 0.101$ ,  $CI = [10.574, 1.067]$ ,  $p_{tukey} < .001$ ). Results held when looking separately within each participant group (Appendix C).

Novelty also had an effect on the number of details recalled per event (Fig. 5b). That is, participants recalled significantly more details for new relative to routine events ( $b = 1.468$ ,  $SE = 0.378$ ,  $CI = [0.542, 2.394]$ ,  $p_{tukey} = .001$ ), and for periodic relative to routine events ( $b = 0.885$ ,  $SE = 0.294$ ,  $CI =$

[0.136, 1.633],  $p_{tukey} = .019$ ). Although they also tended to recall new events in more detail than periodic events, this effect fell short of significance ( $b = 0.583$ ,  $SE = 0.378$ ,  $CI = [-0.178, 1.344]$ ,  $p_{tukey} = .161$ ). The same qualitative pattern was observed in each group separately, although effects were not significant in Group B alone (Appendix C). When interpreting this set of findings, one important consideration is that the number of details recalled per event varied significantly based on the number of details recorded at the time of the diary. As such, it could be that any differences in recall behavior based on event regularity might be driven by how participants describe novel versus familiar experiences *on the day that they occur*, rather than how these different kinds of events are retained in memory across time. To address this possibility, we repeated the same analysis while adding the number of details reported during the diary as an additional covariate. Importantly, in this control model, the effects of event regularity on the number of details recalled remained robust (Appendix C).

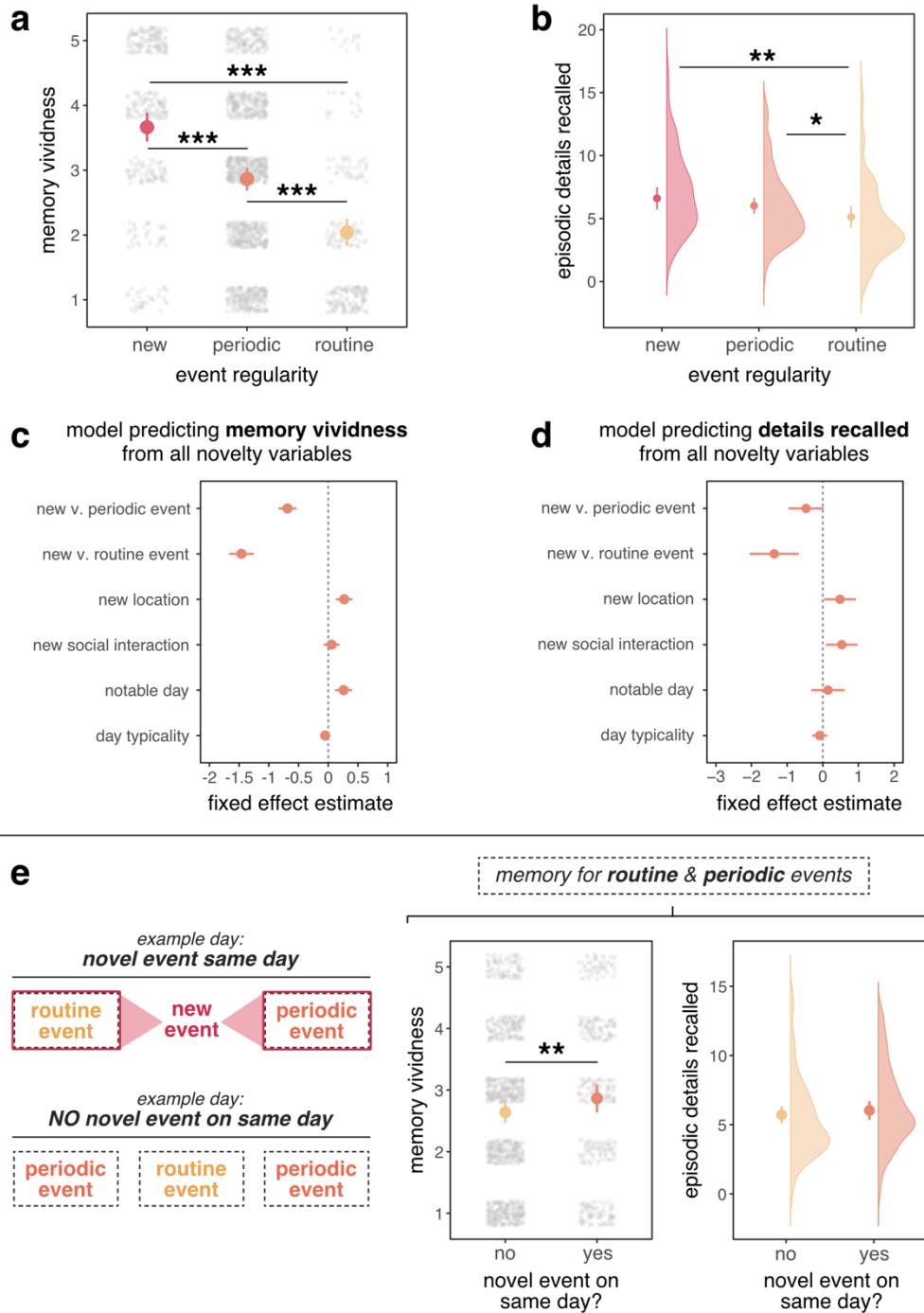
Thus far, we have shown that experiences associated with greater novelty are remembered more vividly and in greater detail, even several weeks after they occur. Our next major aim was to investigate whether novelty had more distributed effects on episodic memory. Specifically, we looked at event memory as a function of whether the participant engaged in any novel or atypical experiences *on the day of the event* — irrespective of whether the event itself was considered novel. In order to characterize the independent influence that each of our novelty measures had on memory, we constructed a single multilevel regression model in which all six measures (i.e., *event regularity*, *day typicality*, *novel locations visited*, and *novel social interactions*, and *day notability*) predicted memory vividness scores (Figure 3.5c). With the exception of event regularity, all of these variables were measured at the granularity of individual days. Due to this model's complexity, we first ran a Frequentist model with only random intercepts (see *Methods*). Here, we saw that memory vividness was significantly higher for events on days where novel locations were visited ( $b = 0.269$ ,  $SE = 0.071$ ,  $CI = [0.128, 0.409]$ ,  $p < 0.001$ ) and on days judged as more notable ( $b = 0.259$ ,  $SE = 0.073$ ,  $CI = [0.116, 0.403]$ ,  $p < 0.001$ ). This model also revealed a significant effect of event regularity, such that new events were remembered more vividly than periodic events, which in turn were remembered more vividly than routine events (new vs. periodic:  $b =$

$0.685$ ,  $SE = 0.077$ ,  $CI = [0.505, 0.866]$ ,  $p_{tukey} < .001$ ; new vs. routine:  $b = 1.462$ ,  $SE = 0.106$ ,  $CI = [1.214, 1.710]$ ,  $p_{tukey} < .001$ ; periodic vs. routine:  $b = 0.777$ ,  $SE = 0.087$ ,  $CI = [0.572, 0.981]$ ,  $p_{tukey} < .001$ ). In contrast, engaging in a novel social interaction did not have a significant effect ( $b = 0.057$ ,  $SE = 0.070$ ,  $CI = [-0.079, 0.193]$ ,  $p = 0.412$ ), nor did ratings of each day's typicality ( $b = -0.051$ ,  $SE = 0.033$ ,  $CI = [-0.114, 0.013]$ ,  $p = 0.116$ ). All of these effects held in a Bayesian regression model with the full random effects structure (Appendix C, Table 5).

We then conducted the same analysis with the number of episodic details as the output variable, again using a model with only random intercepts (Figure 3.5d). Detail counts were significantly higher for events that occurred on a day where a novel location was visited ( $b = 0.483$ ,  $SE = 0.229$ ,  $CI = [0.033, 0.933]$ ,  $p = 0.036$ ) and when a novel social interaction occurred ( $b = 0.532$ ,  $SE = 0.224$ ,  $CI = [0.091, 0.973]$ ,  $p = 0.018$ ). We also saw again that the number of details recalled was greater for new and periodic events relative to routine events (new vs. routine:  $b = 1.364$ ,  $SE = 0.351$ ,  $CI = [0.540, 2.188]$ ,  $p_{tukey} < .001$ ; periodic vs. routine:  $b = 0.895$ ,  $SE = 0.293$ ,  $CI = [0.205, 1.584]$ ,  $p_{tukey} = .007$ ), but did not differ between new and periodic events ( $b = 0.470$ ,  $SE = 0.251$ ,  $CI = [-0.122, 1.061]$ ,  $p_{tukey} = .149$ ). There was no relationship between detail counts at recall and day notability ( $b = 0.145$ ,  $SE = 0.239$ ,  $CI = [-0.324, 0.615]$ ,  $p = 0.544$ ) or day typicality ( $b = -0.087$ ,  $SE = 0.108$ ,  $CI = [-0.299, 0.124]$ ,  $p = 0.417$ ). A corresponding Bayesian model with random slopes for each predictor revealed similar effects, although the effect of visiting a novel location fell short of our criterion for statistical reliability (Appendix C, Table 5). In aggregate, this set of results provides some evidence that in addition to the novelty of the event in question (i.e., whether it was considered “new,” “periodic,” or “routine”), engaging in other novel behaviors on the same day — such as visiting a new location or interacting with somebody new — might also bolster autobiographical event memory.

### ***Evidence for a “novelty penumbra”***

The finding that event memory is enhanced by day-level metrics of novelty hints at the idea that new experiences may exert a penumbra-like effect on autobiographical memory. That is, the act of engaging in something new or unfamiliar might boost memory for other experiences close by in time —



**Figure 3.5. Effects of novelty on event memory.** Subjective memory vividness scores (a) and the number of episodic details recalled (b) for events classified as routine, periodic, and new. Panels (c) and (d) show fixed effect estimates and 95% confidence intervals from a model in which memory vividness scores (c) and the number of episodic details recalled (d) were predicted by all novelty variables (see Methods). Panel (e) compares memory vividness scores (left plot) and episodic details recalled (right plot) for routine and periodic events that either did or did not occur on the same day as a new event. Colored dots and error bars in (a), (b), and (e) indicate estimated marginal means and 95% confidence intervals from multilevel linear regression models. \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

even if those experiences are not novel themselves. In the previous set of analyses, day-level novelty variables did not differentiate between whether a novel experience occurred as part of the recalled event or at some other point in the participant’s day. As such, to more directly explore this idea of a novelty penumbra, we ran an additional analysis in which we focused on memory for events that were explicitly described as “periodic” or “routine”—excluding all “new” events. We then binned these events according to whether or not they occurred *on the same day* as another new event reported by the participant (Figure 3.5e). Our prediction is that non-novel events that happen close in time to new experiences will be better remembered than those that do not. Consistent with this hypothesis, vividness scores were significantly higher for routine and periodic events that occurred on the same day as novel events (relative to those that did not). This effect emerged both in a Frequentist regression model with only random intercepts ( $b = 0.220$ ,  $SE = 0.078$ ,  $CI = [0.067, 0.373]$ ,  $p = 0.005$ ), and a Bayesian model that additionally included random slopes (Appendix C, Table 6). In contrast, the number of details recalled per event did not vary based on the presence of a new event on the same day ( $b = 0.313$ ,  $SE = 0.249$ ,  $CI = [0.249, 0.249]$ ,  $p = 0.210$ ). We also considered the possibility that this novelty penumbra may benefit memory for events on days *adjacent* to the novel experience; however, no support for this relationship was found (Appendix C). Note that we did not run the analyses described in this section separately within each group, given that the number of routine/periodic events that occurred on the same day as novel events was relatively small even in our combined sample.

### 3.3.4 Effects of novelty on temporal memory

Having established that novelty has a beneficial effect on memory for autobiographical events—a benefit that we show extends to non-novel experiences nearby in time—we next turned our attention to temporal memory. In particular, we examined how memory for the order of or time between two experiences might be shaped by the novelty or typicality of what happens *in the intervening time window*. This analysis draws inspiration from previous research on event segmentation, which finds that event boundaries—which can be induced by novel or surprising events (Zacks et al., 2007)—often disrupt memory for the order of events and inflate judgments of temporal distance (DuBrow & Davachi, 2013;

Ezzyat & Davachi, 2011; Heusser et al., 2018). Note that the models in this section included the number of days separating each event pair as a covariate and did not include the small number of event pairs ( $N = 5$  across all participants) that were separated by only one day.

<i>Fixed Effect</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	0.918	0.521	[-0.103, 1.939]	0.078
number of new events	-0.176	0.101	[-0.373, 0.021]	0.080
number of days with new locations visited	-0.000	0.241	[-0.473, 0.473]	1.000
number of days with novel social interactions	-0.001	0.229	[-0.450, 0.447]	0.996
mean day typicality	-0.094	0.106	[-0.301, 0.114]	0.376
number of notable days	0.209	0.156	[-0.096, 0.515]	0.180
number of days between pair	-0.004	0.176	[-0.349, 0.340]	0.980

**Table 3.1.** Results from a multilevel logistic regression model in which temporal order accuracy for each event pair was predicted by novelty experienced during the days between the two events.

Mirroring our approach for event memory analyses, we constructed a logistic regression model in which order accuracy was predicted by all of our novelty constructs. Critically, these novelty variables were extracted from daily diary data for the days *between* each event pair (i.e., “intervening days”). The full set of predictors included: the number of “new” events on intervening days, the number of intervening days with novel locations visited, the number of intervening days with novel social interactions, the number of intervening days deemed to be “notable,” and the average typicality of intervening days. This model revealed a marginal effect of the number of new events, such that the likelihood of correctly recounting the order of two events was slightly lower when there were more new experiences in between them ( $b = -0.176$ ,  $SE = 0.101$ ,  $CI = [-0.373, 0.021]$ ,  $p = 0.080$ ). Beyond this, none of our additional novelty measures were significantly associated with temporal order accuracy (Table 3.1), and these null results held within each group (Appendix C). We then repeated the same analysis with subjective temporal distance as our outcome variable. Here we found a marginal relationship between distance ratings and the number of intervening novel events, such that event pairs separated by more new events were remembered as being closer together in time ( $b = -0.077$ ,  $SE = 0.046$ ,  $CI = [-0.167, 0.012]$ ,  $p = 0.091$ ) — an effect that was driven by participants in Group A (Appendix C). However, no other

significant associations emerged (Table 3.2). Taken together, although we observed some weak evidence that memory for the temporal relationship between two events is affected by the presence of novel experiences between them, such findings were not nearly as striking as those of novelty on memory for individual events.

<i>Fixed Effect</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>95% CI</i>	<i>p</i>
(Intercept)	2.281	0.231	[1.826, 2.735]	< 0.001
number of new events	-0.077	0.046	[-0.167, 0.012]	0.091
number of days with new locations visited	0.071	0.109	[-0.143, 0.284]	0.516
number of days with novel social interactions	0.070	0.106	[-0.138, 0.277]	0.511
mean day typicality	0.011	0.048	[-0.084, 0.106]	0.820
number of notable days	0.024	0.070	[-0.114, 0.162]	0.734
number of days between pair	0.187	0.075	[0.039, 0.336]	0.013

**Table 3.2.** Results from a multilevel linear regression model in which subjective temporal distance ratings for each event pair were predicted by novelty experienced during the days between the two events.

### 3.4 Discussion

In this study, we found that novel autobiographical events were remembered more vividly and in greater detail than those perceived as more typical. The mnemonic benefit of engaging in a new experience also extended across time, such that other events reported on the same day as a novel event — or on the same day as visiting a novel location or interacting with a new person — were also remembered with greater vividness or detail. In contrast to these effects, the impact of novelty on temporal memory was minimal. These findings echo recent studies of autobiographical memory indicating memory enhancements when novel people or places are involved (Bainbridge & Baker, 2022) or when events occur amidst a surprising world event (e.g., the onset of COVID-19, Rouhani et al., 2023).

Novel events may be better remembered through several mechanisms, such as by recruiting greater attentional resources during encoding or providing more distinctive cues during memory retrieval. In our dataset, we found that new or atypical experiences on a given day also affect memory for other, non-novel events occurring on that same day. These effects are consistent with previous evidence that novelty can exert a penumbra-like effect on memory, such that new experiences facilitate the encoding

(Ballarini et al., 2009, 2013; Li et al., 2003; Moncada & Viola, 2007; but see Biel & Bunzeck, 2019; Quent & Henson, 2022) and retrieval (Duncan et al., 2012) of other events nearby in time. As mentioned in the *Introduction*, most findings suggest that this benefit arises only when novelty and encoding occur in close temporal proximity (i.e., within the span of an hour or two). This limit theoretically stems from the neurobiological process by which the novelty penumbra is thought to occur, referred to as “synaptic tag-and-capture.” In this process, a weak learning experience creates a synaptic “tag” on the newly-formed memory trace, which is then potentiated by the release of plasticity-related proteins that is triggered by a strong learning experience — such as novelty, arousal, or emotion (Dunsmoor et al., 2015, 2022; Moncada et al., 2015; Redondo & Morris, 2011). Because this synaptic tag is transient in nature, it is thought that these “strong” and “weak” learning experiences must occur within a critical time window. In the current study, although we did not collect precise information about the timing of each reported event, it is possible (and perhaps likely) that novel events did not always occur in close proximity to the routine and periodic events that incurred memory benefits. One explanation, consistent with some previous findings (Ramirez Butavand et al., 2020), is that the reach of a novel experience can sometimes extend across many hours.

Another interesting possibility, however, is that our effects are driven by the act of completing a diary entry at the end of each day. To some extent, describing events as part of this diary can be thought of as a conscious and deliberate form of post-encoding memory reactivation. A long history of prior work speaks to the functional significance of memory reactivation, finding that it helps to strengthen and consolidate memories (Gruber et al., 2016; Hermans et al., 2017; Hu et al., 2020; Schapiro et al., 2018; Tambini & Davachi, 2013, 2019), as well as to integrate new memories with existing knowledge structures (Morton et al., 2023; Schlichting & Preston, 2014; van Kesteren et al., 2020). Reactivation can also push memory traces into a more labile state, allowing them to be more easily updated or transformed (De Oliveira Alvares et al., 2013; Jardine et al., 2022; Kuhl et al., 2012; Nader, 2003; St. Jacques & Schacter, 2013). In light of these ideas, it could be that reactivating multiple memories in close succession during each daily diary allows the salience ascribed to a novel experience to “spread” to other

simultaneously active memories. Other researchers have pointed out that the deliberate recording of experiences involved in diary-style paradigms might introduce unwanted confounds to any investigation of naturalistic memory function (e.g., Brewer, 1988; Finley & Brewer, 2024). While this is a reasonable point, it is also true that reflecting on your day as it draws to a close — either by yourself or with others — is not an unusual activity. Moreover, changing how a person reactivates their recent memories, rather than asking them to alter aspects of their actual lives, is arguably a more feasible target for therapeutic interventions. Indeed, the “HippoCamera” smartphone app was explicitly designed to promote the conscious reactivation of memories, and has been shown to improve autobiographical memory in a sample of older adults (Martin et al., 2022).

In our study, we focused on two measures of episodic event memory: subjective vividness ratings and the number of episodic details recalled. These measures were highly correlated and showed similar effects throughout many of our analyses. That being said, associations between novelty and event memory were generally stronger when looking at subjective vividness ratings, and only vividness showed a significant novelty penumbra effect. There are several potential reasons for divergence between these measures. The first draws from previous work suggesting that participants’ metacognitive judgments about the strength of a memory do not always correspond with the accuracy of what is remembered. This is especially true in the case of experiences that are emotional or unexpected. For example, “flashbulb memories” refer to memories for events that are highly surprising or emotionally charged (R. Brown & Kulik, 1977; Hirst & Phelps, 2016; Winograd & Neisser, 2006). Although people tend to recall these experiences with strong confidence, even after substantial time has passed, flashbulb memories are not exceptionally accurate (Chiew et al., 2022; Hirst et al., 2015; Talarico & Rubin, 2003; Wright, 1993). Given this dissociation between objective and subjective memory outcomes, it could be that novelty exerts a greater effect on individuals’ “feeling of remembering” than it does on how much they actually recall. However, it is also important to note that our analyses involving the number of episodic details recalled had less statistical power than those using vividness, given that only a third of reported events were included in the free recall test, whereas vividness ratings were available for all events (in both the

event recall test and the temporal memory test). There is also existing evidence that for young adults, vividness ratings are typically a reliable index of memory content (Folville et al., 2022). Additional research is therefore needed to resolve the extent to which novelty's effects on autobiographical memory are driven subjective feelings of remembering versus objective memory richness.

As noted above, we did not observe any significant effects of novelty on temporal memory performance. This lack of results is perhaps surprising in light of prior research on how temporal memory is shaped by contextual stability and change. Based on models of event segmentation, novel events can create “boundaries” in a person’s ongoing experience, either by triggering prediction errors or by causing salient changes in one’s environment or mental state (Clewett et al., 2019; Clewett & Davachi, 2017; Shin & DuBrow, 2021; Zacks et al., 2007). In the Ezzyat-DuBrow-Davachi (EDD) paradigm (Buonomano et al., 2023), boundaries elicit remarkably reliable effects on temporal memory: namely, memory for the order of two items is disrupted when those items cross a boundary, and people tend to remember items crossing a boundary as occurring further apart in time (Clewett, Gasser, et al., 2020; DuBrow & Davachi, 2013; Ezzyat & Davachi, 2011, 2014; Heusser et al., 2018). Extrapolating from these findings, one might predict that if a participant reports more novel experiences within a given time period, they will show reduced memory for the order of events spanning that period, as well as inflated judgments of temporal distance between those events. In this study, we observed a trend toward the former prediction, such that order memory was marginally worse for event pairs that spanned days where more new events were reported. However, distance judgments for these pairs trended toward greater temporal *compression* rather than inflation. Although marginal effects should be interpreted with caution, one possibility is that the effects of novelty and temporal memory might depend on the strategy used to judge temporal relationships. For example, a person could remember which of two events occurred earlier in time by explicitly recalling the “chain” of experiences connecting them between them (Janssen et al., 2006; Lewandowsky & Murdock Jr., 1989; Logan & Cox, 2021)—which might be easier when more novelty occurred—or by relying on a more general sense of which event “feels” more remote. Which strategy is adopted might depend on many factors, such as the actual amount of time separating two events or how

remote events are to the present (DuBrow & Davachi, 2017; Janssen et al., 2006). Future work should explore these ideas by testing temporal memory at multiple delays for events that span a wider range of temporal distances, while also querying participants' use of different cognitive strategies.

There are several methodological limitations in the present study worth considering. First, while we interpret the number of details recalled per event as a relatively objective metric of memory recall (compared to memory vividness ratings), we did not explicitly evaluate each detail for its accuracy. Although accuracy could be approximated by comparing the details recalled at test with event descriptions from the diary period, it is possible that any details present at recall but not in the original diary description reflect correctly remembered information that was simply not recorded during the diary period. As such, following previous studies of autobiographical memory (e.g., Folville et al., 2023; Levine et al., 2002; Rouhani et al., 2023), we chose to assume that the information generated during recall was generally correct. For similar reasons, we also did not account for the fact that the event titles used to cue participants' memory sometimes "gave away" information about the event in question. Because details revealed by event titles might nevertheless still reflect real memories, and because the number of details recalled per event was relatively small to begin with (~6 details on average), we included all details in our analyses. Despite these caveats, however, the diary approach employed here still has significant advantages over other methods commonly used to investigate autobiographical memory, such as asking participants to retroactively recall experiences that took place weeks, months, or years in the past (e.g., Addis et al., 2004; Bonnici et al., 2012; Crovitz & Schiffman, 1974; Levine et al., 2002; Robinson, 1976). Relative to the current study, these other paradigms are more limited to examining memory only for very meaningful or salient events, as these are the memories that come to mind most readily (without any targeted cueing). Daily diaries also give us access to a rich swath of contextual information about what participants were doing around the time when each memory was formed — which is typically unavailable in purely retrospective reports.

It is also important to consider the approach we used to quantify novelty. As described throughout the manuscript, we used an assortment of different variables collected during the diary period for this

purpose, allowing us to approximate the degree to which participants engaged in new or atypical experiences throughout the study. However, these measures do not provide a complete picture of all the new experiences that each participant had, nor do they allow us to draw precise conclusions about exactly what kinds of novelty were experienced. Previous research has perhaps focused most heavily on the memory benefits of exploring a novel spatial environment (Li et al., 2003; Schomaker, 2019; Schomaker et al., 2014). Consistent with this literature, we found that even after accounting for the novelty of an event itself, visiting a new location enhanced memory vividness for events occurring on the same day. We also found that the number of details recalled about an event was higher when the event occurred on a day where a novel social interaction occurred. To our knowledge, there are very few previous studies that have directly tested how interacting with new people affects episodic memory, aside from some evidence that videos including unfamiliar people were better remembered than those including familiar people (Bainbridge & Baker, 2022). One intriguing follow-up question to our findings is whether different kinds of novel experiences impact memory in divergent ways. For example, an event occurring in a new location might orient a person's attention to perceptual information, whereas new social interactions might lead them to prioritize the emotional aspects of an experience. Future research in this space should collect more fine-grained information about the multitude of features that could characterize an event as novel, as well as how these features modulate different aspects of episodic memory.

In sum, here we used an intensive longitudinal daily diary study to examine the relationship between experiential novelty and real-world autobiographical event memory. Not only were novel events remembered more vividly and in greater detail than those perceived as more typical, they also helped bolster memory for non-novel experiences close in time. This study builds upon other findings benefits of novelty and cognitive and affective function. Novelty and variety in one's experience have been associated with enhanced well-being (Folville et al., 2023; Heller et al., 2020; S. Lee et al., 2016), with some researchers arguing that novelty satisfaction may be a basic psychological need (Bagheri & Milyavskaya, 2020; González-Cutre et al., 2020). Going forward, additional longitudinal studies with larger samples will be necessary to validate and extend our findings, addressing questions such as which

forms of novelty have the most potent effects on episodic memory, as well as in what circumstances autobiographical experiences benefit instead from greater predictability and familiarity. Ultimately, our findings affirm the importance of examining memory function in real-world settings, where what we remember is heavily shaped by the structure and content of our everyday lives.

## Conclusion

In this dissertation, I have presented a series of behavioral and neuroimaging experiments that explore how the predictability and novelty embedded within our experience impact episodic memory. Such work speaks to the flexible and dynamic nature of the human memory system, highlighting that what we remember depends on the context in which it is learned — a context that invariably includes our knowledge and expectations.

Chapter 1 introduced an original behavioral paradigm to investigate how predictable motor behavior affects simultaneous episodic encoding. In designing this study, we drew inspiration from the fact that motor actions are central to how humans navigate the world, and many of our behaviors are highly familiar and repetitive. Yet, the effects of familiar behavior on novel learning have rarely been explored in prior research. Across three experiments, we found that executing a predictable (versus random/variable) action sequence during the encoding of novel items enhanced memory for the temporal order of those items. This benefit was largely selective to order memory, suggesting that was not simply a reflection of increased attention at encoding. Instead, we argue that the predictable action sequence functioned as a kind of organizational “scaffold” for incoming stimuli, and that the mental maintenance of this scaffold throughout each encoding event facilitated temporal integration between successive items.

In Chapter 2, we sought further evidence for this scaffolding mechanism by using fMRI to probe how predictable motor sequence knowledge influences brain activity during encoding. First, we replicated the behavioral finding from Chapter 1 that the execution of a predictable action sequence facilitates order memory for concurrent item sequences. Multivariate fMRI analyses then revealed that the predictable action sequence was represented within the default mode network — i.e., the medial prefrontal cortex (mPFC) and posterior medial cortex (PMC) — as well as in premotor cortex. Critically, stronger representations of this “action scaffold” in PMC and premotor cortex during encoding were associated with more stable activity across time in the hippocampus and lateral occipital cortices. Greater neural stability in the hippocampus (as well as in other neocortical regions), in turn, was linked to better

subsequent order memory. Taken together, these findings shed light on how the brain encodes the temporal structure of episodic events, particularly when those events involve predictable sequential behavior. They also reveal how cortical representations of sequential knowledge might modulate encoding-related activity in regions that process incoming stimuli (e.g., the hippocampus).

Finally, in Chapter 3, we moved beyond motor sequence knowledge to ask how autobiographical event memory is impacted by novelty versus typicality in everyday life. To this end, we asked participants to complete “daily diary” entries every day for two weeks, allowing us to probe memory for their personal experiences. These diaries also provided us with information about the extent to which participants engaged in novel activities each day, such as visiting a new location or interacting with an unfamiliar person. Contrary to our predictions, we found that novelty embedded in daily experience had no clear effects on participants’ memory for the sequential order of or temporal distance between autobiographical events. However, novel events were remembered more vividly and with greater detail than events that were considered typical. Novelty also demonstrated a penumbra-like effect on event memory, such that non-novel events recorded on the same day as a novel experience were also remembered more vividly.

Taken together, these findings reveal multiple ways in which both predictability and novelty impact what we learn and remember. They also add to our knowledge of the factors that influence memory for how experiences are structured across time — an intrinsic feature of episodic memory organization that is not always captured by experimental paradigms. Below, I discuss several open questions related to this research that future work could pursue.

### **When does predictability help versus hurt memory?**

If both consistency and inconsistency with prior knowledge can enhance episodic memory, a natural follow-up question is: when is predictable structure beneficial and when is it detrimental? In the context of this dissertation, part of the answer to this question can be resolved by considering methodological differences in the paradigms used for Chapters 1 and 2 versus Chapter 3. In Chapters 1 and 2, the predictable element of participants’ experience was the learned motor action sequence they

executed during events. We then probed their memory for novel items encoded *in parallel* to the execution of that action sequence. As such, although participants' experience in the predictable store was consistent with prior knowledge, the actual items they encoded were not known in advance (beyond their semantic category). In Chapter 3, we showed that memory was enhanced for autobiographical events that were themselves defined as novel, whereas memory for routine events was weaker. Drawing from these results, it is likely that an event must contain *some* meaningful novel information in order for episodic memory to benefit from predictable structure — a quality that may not apply to events that participants define as routine. After all, if there is nothing sufficiently new about an experience, there is no adaptive reason to store it in memory.

Other factors likely to mediate the effects of predictability on episodic memory are one's goals or motivations during learning. In Chapters 1 and 2, participants knew throughout the experiment that their task was to remember the sequence of items collected during each errand. They were therefore incentivized to keep their attention focused on item stimuli throughout events, even when aspects of their behavior (i.e., the well-learned action sequence) could be enacted without much thought. In everyday life, this demand for encoding is not always present in familiar contexts. If a person takes the same subway route every day for work, this knowledge *could* scaffold memory for whatever novel experiences occur in parallel, but it could also permit the individual to tune out their surroundings, letting their mind wander as they progress through their commute on autopilot. Such lapses in attention, in turn, are likely to disrupt episodic memory encoding (deBettencourt et al., 2018; Jayakumar et al., 2023). Predictability can also impair cognition even when the motivation to attend to one's surroundings is high. In the context of driving, for example — where the consequences of zoning out are potentially catastrophic — people are still prone to attentional lapses when taking familiar routes (Burdett et al., 2018; Charlton & Starkey, 2013; Harms, 2023). Ultimately, the ability of predictable structure to scaffold memory is likely contingent on numerous interrelated factors, including a person's current goals, their attentional bandwidth, and how fully automated their behavior has become.

Other research has sought to explain conflicting findings about when prior knowledge helps

versus hurts new learning by considering the relationship between one's prior knowledge and what is being learned. For example, some work has differentiated between "knowledge relevance" and "knowledge congruence" (Brod & Shing, 2018): knowledge relevance refers to whether new information can reasonably be linked to a pre-existing memory network, whereas congruence refers to the subjective sense of "fit" between prior knowledge and new information. In this study, while both knowledge relevance and congruence had beneficial effects on associative memory, encoding activity in vmPFC only tracked congruence. In a similar vein, another recent study independently varied the extent to which a novel event was consistent with schematic knowledge versus whether it was actually predicted by a participant — finding that these two features had dissociable but interactive effects on memory (Huang et al., 2023). In the experimental paradigm we used for Chapters 1 and 2, the novel item sequences encoded in the predictable context had no meaningful relationship to the learned action sequence, whereas in Chapter 3, the qualities that drove participants to identify their experiences as novel versus routine were not explicitly recorded. As such, while it is difficult within the current body of work to disentangle what aspects of the relationship between prior knowledge and new learning lead to facilitation versus hindrance, this is an intriguing avenue for future research.

Finally, the extent to which predictable structure aids memory likely depends on how and when memory is tested. As noted in the introduction, although prior knowledge often facilitates memory for related information, it can also increase susceptibility to memory errors (Kleider et al., 2008; Lew & Howe, 2017; Tompary & Thompson-Schill, 2021; Webb et al., 2016). In these circumstances, an individual might accurately remember the gist of a past experience, but be unable to discriminate between details that actually occurred versus those that seem plausible based on past experience (Loftus et al., 1978; Pardilla-Delgado & Payne, 2017; Roediger & McDermott, 1995). When a task requires memory to be highly precise, therefore, greater activation of prior knowledge during learning might impair performance rather than facilitate it. The experiments used throughout this dissertation also varied in the time delay between encoding and retrieval: whereas Chapters 1 and 2 probed temporal memory within a few minutes of each encoding event, the daily diary paradigm in Chapter 3 asked participants to recall

memories that were multiple weeks old. It is possible that novelty may be particularly beneficial for episodic memory when memories must be retained across long delays.

In sum, this discussion highlights several sources of complexity that should be considered in future explorations of how predictability and novelty influence memory, including an individual's goals, the extent to which predictability elicits attention lapses, the relationship between prior knowledge and new information, and the specificity and timing of memory retrieval.

## **Effects across the lifespan of a memory**

Episodic memory processing operates in multiple stages, starting with the initial *encoding* of an experience, followed by *consolidation* processes that stabilize or transform the memory trace, and then at last by memory *retrieval*. These stages reflect multiple timepoints at which the effects of prior knowledge and/or novelty on memory could emerge. Throughout this dissertation, I have most heavily focused on how these factors might shape memories at the time of encoding. Activation of prior knowledge during encoding theoretically allows novel memory traces to become linked to existing knowledge networks, enhancing memory strength or stability (Gilboa & Marlatte, 2017; McClelland, 2013). The neuroimaging results described in Chapter 2 are broadly consistent with this idea, showing that activation of motor sequence knowledge (and/or other relevant contextual features) during encoding was linked to more stable brain activity and better temporal order memory. However, previous research also indicates that prior knowledge can modulate memory processing after learning, such as by altering neural signatures of post-encoding consolidation (Coutanche & Thompson-Schill, 2014; Schlichting & Preston, 2016; Tse et al., 2007) or increasing consolidation-related gains in memory performance after sleep (Henderson & James, 2018; King et al., 2019; Zion et al., 2019). At retrieval, in turn, prior knowledge might guide access to targeted episodic content. Indeed, retrieval of knowledge-consistent information has been associated with increased activity in mPFC and other default mode regions (Brod et al., 2015; Guo et al., 2023; van Buuren et al., 2014; van Kesteren, Rijpkema, et al., 2010) and modified patterns of hippocampal-mPFC connectivity (Bonasia et al., 2018; Guo et al., 2023; Guo & Yang, 2020). In our

“errand running” paradigm, therefore, it is possible that the predictable action scaffold also exerted effects on memory-related processing in the short interval between encoding and retrieval and/or during the order memory test. Future analyses of this fMRI dataset will consider both of these possibilities.

Like prior knowledge, novelty can also influence the strength of a memory at multiple stages in its lifespan. Novelty detection tends to sharpen or narrow attention to relevant stimuli and increase an agent’s motivation to learn, boosting the effectiveness of encoding (Düzel et al., 2010; Kafkas & Montaldi, 2018). It can also shape mechanisms of memory consolidation, such as by facilitating long-term potentiation (Davis et al., 2004; Lisman & Grace, 2005) or by enhancing cross-regional coupling within memory networks (Cowan et al., 2021). Even during retrieval, encountering novel information can improve discrimination between previously-encoded stimuli and similar lures (Duncan et al., 2012). In Chapter 3 of this dissertation, we introduced two hypotheses to explain how novelty benefits autobiographical event memory — particularly for non-novel events from the same day. The first explanation pointed to a “tag-and-capture” mechanism, arguing that the weak memory trace created by a non-novel event might have been potentiated by plasticity-related proteins released in response to the novel experience (Dunsmoor et al., 2022; Lorents et al., 2023; Redondo & Morris, 2011; Schomaker, 2019). The second hypothesis was tied more directly to the design of our study, where participants recorded information about three distinct episodic memories as part the daily diaries completed each night. The process of reactivating multiple memories together in this manner may have allowed the motivational significance associated with a new event to “spread” to memories for non-novel experiences. Indeed, this mechanism is similar to that proposed by Clewett et al. (2022) to explain how an aversive experience can rescue memories for neutral events encoded close by in time. In practice, clarifying the neurological basis of the novelty-induced benefits in Chapter 3 would be difficult, given the highly naturalistic design of the study. However, immersive encoding paradigms — like those that make use of virtual reality (VR) technology — might prove useful in clarifying whether or not post-learning reactivation helps to broaden the mnemonic benefits of novelty.

## **Remembering the temporal structure of experience**

Throughout this dissertation, I have highlighted the role that contextual information plays in how we remember temporal structure. Such ideas build upon two related theoretical frameworks — temporal context models and event segmentation theory — which suggest that episodic memories are formed on a backdrop of drifting temporal context representations, which shift rapidly at event boundaries (i.e., the transition from one context to another) (Clewett et al., 2019; DuBrow et al., 2017; Zacks, 2020). As a result of this process, stimuli encountered close together in time (and that belong to the same event context) tend to be grouped together in memory. The results of Chapters 1 and 2 suggest that prior knowledge about sequential structure might modulate this mechanism by boosting the stability or consistency of context representations within an event, increasing the strength with which constituent items become integrated in memory.

How context stability influences temporal memory organization at larger scales — when memories are separated by days or weeks instead of minutes or hours — is less clear. One recent study tackled this limitation by testing memory for events from popular, multi-season TV shows (e.g., Game of Thrones), where episodes are often watched across multiple days, months, or years (Yousif et al., 2024). Interestingly, this study revealed several effects that appear in conflict with what is found in paradigms using shorter timescales: for example, order memory for the TV show events was generally worse when they occurred *within* the same “context” (i.e., episode or season) versus different contexts. In experiments looking at shorter timescales, order memory for items from the same event is typically *better* than for item pairs that cross a boundary, and within-event pairs are also remembered as being separated by smaller temporal distances (Clewett et al., 2019; DuBrow & Davachi, 2014; Heusser et al., 2018). In Chapter 3 of this work, we found only marginal evidence that novelty impacted temporal memory: specifically, autobiographical events separated by more novel events — which arguably create more boundary-like divisions in experience — were slightly less likely to be remembered in the correct order, but were also remembered as occurring slightly further apart in time. While the former conforms with findings from

short-timescale experiments, the latter does not. In light of this mixed evidence, one worthwhile avenue for future research is clarifying what factors shape the temporal organization of memories separated by long time intervals. For example, one might consider the importance of the relationship between the events: are they thematically connected like episodes in a TV series, or are they completely unrelated? It is also worth noting that in our exploration of how prior knowledge and novelty shape temporal memory, we have focused mostly on memory for sequential order (with Chapter 3 also probing memory for subjective temporal distance). However, one could also test how these factors influence memory for other kinds of temporal information — such as how much time elapsed within an event, or when exactly a specific experience occurred.

## **Implications for memory interventions**

One of the long-term goals implicit in memory research is to understand how deficits in memory function — whether due to neurological disease or healthy aging — can be remedied. Although the aims of this dissertation were not clinical or applied in nature, our results hint at several strategies that may eventually prove useful for mitigating episodic memory dysfunction. In particular, our findings that familiar motor behavior can scaffold episodic encoding highlight the possibility of leveraging one memory system to compensate for another. Individuals with Alzheimer's disease, for example, are characterized by profound deficits in episodic memory but relatively intact functioning in other memory systems (e.g., skill learning; Eslinger & Damasio, 1986; van Halteren-van Tilborg et al., 2007). It is thus worth exploring whether executing familiar motor behaviors during encoding can compensate for this selective dysfunction. Related to this idea, there is some evidence that listening to music — which often has familiar or predictable structure — can improve episodic memory in Alzheimer's patients (Irish et al., 2006; Simmons-Stern et al., 2010; although it should be noted that other mechanisms, such as music's effects on mood or arousal, might explain this effect). In addition, further research on the interactions between motor behavior and episodic memory might help clarify the basis of memory deficits in patients with impaired motor function, such as what occurs in Parkinson's disease (Breen & Drutte, 2013; Chung

et al., 2021; Shohamy et al., 2005; Siquier & Andrés, 2021).

As mentioned briefly in Chapter 3, another potential application of this research concerns the use of daily diary-style interventions to improve autobiographical memory. Other work has introduced multiple tools — such as SenseCam and the HippoCamera — that record moments from an individual's lived experience so they can be reviewed later on, finding that use of these tools can boost memory recall in aging and disease populations (Berry et al., 2007; Chow & Rissman, 2017; Dubourg et al., 2016; Martin et al., 2022). Although camera-based tools are likely more effective than written diary entries at enhancing memory (Dubourg et al., 2016), our findings raise the possibility that manipulating the novelty of experiences might enhance the efficacy of such interventions. For example, when a person reviews pictures or videos of their recent experiences, ensuring that at least one new or atypical event is included as part of each "review session" might strengthen retention of all revisited memories.

## Final remarks

Ultimately, understanding how prior knowledge and novelty influence what we remember is a complex undertaking, raising numerous questions about the type of memory to be tested, the kind of knowledge or novelty under consideration, and the neural mechanisms by which such influences arise. With the body of work presented here, I aimed to chip away at this constellation of questions, making use of diverse experimental paradigms, analytical techniques, and both behavioral and neuroimaging data. Across three chapters, I showed that: (1) well-learned motor behavior can scaffold memory for novel episodic events; (2) that this memory facilitation is associated with the neural representation of motor knowledge during encoding and the stability of brain activity across time; and (3) that everyday novel experiences are not only better remembered than typical ones, but can also enhance memory for other events happening that same day. Together, these effects reveal several compelling ways in which our cognition is colored by the knowledge we carry with us.

## References

- Åberg, C. S., & Nilsson, L.-G. (2003). A strict response criterion yields a mirror effect in the novelty paradigm. *Scandinavian Journal of Psychology*, 44(5), 425–432. <https://doi.org/10.1046/j.1467-9450.2003.00363.x>
- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8. <https://doi.org/10.3389/fninf.2014.00014>
- Addis, D. R., Moscovitch, M., Crawley, A. P., & McAndrews, M. P. (2004). Recollective qualities modulate hippocampal activation during autobiographical memory retrieval. *Hippocampus*, 14(6), 752–762. <https://doi.org/10.1002/hipo.10215>
- Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, 93(2), 203. <https://doi.org/10.1037/0033-2909.93.2.203>
- Albouy, G., Fogel, S., King, B. R., Laventure, S., Benali, H., Karni, A., Carrier, J., Robertson, E. M., & Doyon, J. (2015). Maintaining vs. enhancing motor sequence memories: Respective roles of striatal and hippocampal systems. *NeuroImage*, 108, 423–434. <https://doi.org/10.1016/j.neuroimage.2014.12.049>
- Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., Darsaud, A., Ruby, P., Luppi, P.-H., Degueldre, C., Peigneux, P., Luxen, A., & Maquet, P. (2008). Both the Hippocampus and Striatum Are Involved in Consolidation of Motor Sequence Memory. *Neuron*, 58(2), 261–272. <https://doi.org/10.1016/j.neuron.2008.02.008>
- Anderson, J. R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition*, 9(3), 237–246. <https://doi.org/10.1111/mbe.12110>
- Anderson, R. C., Spiro, R. J., & Anderson, M. C. (1978). Schemata as Scaffolding for the Representation of Information in Connected Discourse. *American Educational Research*, 15(3), 433–440.
- Audrain, S., & McAndrews, M. P. (2022). Schemas provide a scaffold for neocortical integration of new

- memories over time. *Nature Communications*, 13(1), Article 1. <https://doi.org/10.1038/s41467-022-33517-0>
- Avants, B., Tustison, N. J., & Song, G. (2009). Advanced Normalization Tools: V1.0. *The Insight Journal*. <https://doi.org/10.54294/uvnhin>
- Bagheri, L., & Milyavskaya, M. (2020). Novelty–variety as a candidate basic psychological need: New evidence across three studies. *Motivation and Emotion*, 44(1), 32–53.  
<https://doi.org/10.1007/s11031-019-09807-4>
- Bainbridge, W. A., & Baker, C. I. (2022). Multidimensional memory topography in the medial parietal cortex identified from neuroimaging of thousands of daily memory videos. *Nature Communications*, 13(1), 6508. <https://doi.org/10.1038/s41467-022-34075-1>
- Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron*, 95(3), 709-721.e5.  
<https://doi.org/10.1016/j.neuron.2017.06.041>
- Baldassano, C., Hasson, U., & Norman, K. A. (2018). Representation of Real-World Event Schemas during Narrative Perception. *The Journal of Neuroscience*, 38(45), 9689–9699.  
<https://doi.org/10.1523/JNEUROSCI.0251-18.2018>
- Ballarini, F., Martínez, M. C., Perez, M. D., Moncada, D., & Viola, H. (2013). Memory in Elementary School Children Is Improved by an Unrelated Novel Experience. *PLOS ONE*, 8(6), e66875.  
<https://doi.org/10.1371/journal.pone.0066875>
- Ballarini, F., Moncada, D., Martinez, M. C., Alen, N., & Viola, H. (2009). Behavioral tagging is a general mechanism of long-term memory formation. *Proceedings of the National Academy of Sciences*, 106(34), 14599–14604. <https://doi.org/10.1073/pnas.0907078106>
- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11(7), 280–289. <https://doi.org/10.1016/j.tics.2007.05.005>
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology* (pp. xix, 317). Cambridge University Press.

- Bass, W., & Oswald, K. (2014). Proactive control of proactive interference using the method of loci. *Advances in Cognitive Psychology*, 10(2), 49–58. <https://doi.org/10.5709/acp-0156-3>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015a). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015b). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Bein, O., & Davachi, L. (2024). Event Integration and Temporal Differentiation: How Hierarchical Knowledge Emerges in Hippocampal Subfields through Learning. *The Journal of Neuroscience*, 44(10), e0627232023. <https://doi.org/10.1523/JNEUROSCI.0627-23.2023>
- Bein, O., Plotkin, N. A., & Davachi, L. (2021). Mnemonic prediction errors promote detailed memories. *Learning & Memory*, 28(11), 422–434. <https://doi.org/10.1101/lm.053410.121>
- Bein, O., Reggev, N., & Maril, A. (2014). Prior knowledge influences on hippocampus and medial prefrontal cortex interactions in subsequent memory. *Neuropsychologia*, 64, 320–330. <https://doi.org/10.1016/j.neuropsychologia.2014.09.046>
- Bein, O., Reggev, N., & Maril, A. (2020). Prior knowledge promotes hippocampal separation but cortical assimilation in the left inferior frontal gyrus. *Nature Communications*, 11(1), Article 1. <https://doi.org/10.1038/s41467-020-18364-1>
- Bein, O., Trzewik, M., & Maril, A. (2019). The role of prior knowledge in incremental associative learning: An empirical and computational approach. *Journal of Memory and Language*, 107(November 2017), 1–24. <https://doi.org/10.1016/j.jml.2019.03.006>
- Bellana, B., Mansour, R., Ladyka-Wojcik, N., Grady, C. L., & Moscovitch, M. (2021). The influence of prior knowledge on the formation of detailed and durable memories. *Journal of Memory and Language*, 121, 104264. <https://doi.org/10.1016/j.jml.2021.104264>

- Bellmund, J. L. S., Deuker, L., & Doeller, C. F. (2019). Mapping sequence structure in the human lateral Entorhinal cortex. *eLife*, 8, 1–20. <https://doi.org/10.7554/eLife.45333>
- 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), 3395. <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), 1–13. [https://doi.org/10.1162/jocn\\_a\\_01592](https://doi.org/10.1162/jocn_a_01592)
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.
- Benjamini, Y., Yekutieli, D., Edwards, D., Shaffer, J. P., Tamhane, A. C., Westfall, P. H., & Holland, B. (2005). False Discovery Rate: Adjusted Multiple Confidence Intervals for Selected Parameters. *Journal of the American Statistical Association*, 100(469), 71–93.
- Berlyne, D. E. (1951). Attention to change. *British Journal of Psychology*, 42, 269–278.
- Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B., & Wood, K. (2007). The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: A preliminary report. *Neuropsychological Rehabilitation*, 17(4–5), 582–601.  
<https://doi.org/10.1080/09602010601029780>
- Biel, D., & Bunzeck, N. (2019). Novelty Before or After Word Learning Does Not Affect Subsequent Memory Performance. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01379>
- Blokland, A. (1995). Acetylcholine: A neurotransmitter for learning and memory? *Brain Research. Brain Research Reviews*, 21(3), 285–300. [https://doi.org/10.1016/0165-0173\(95\)00016-x](https://doi.org/10.1016/0165-0173(95)00016-x)
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary Methods: Capturing Life as it is Lived. *Annual Review of Psychology*, 54, 579–616. <https://doi.org/10.1146/annurev.psych.54.101601.145030>
- Bonasia, K., Sekeres, M. J., Gilboa, A., Grady, C. L., Winocur, G., & Moscovitch, M. (2018). Prior

knowledge modulates the neural substrates of encoding and retrieving naturalistic events at short and long delays. *Neurobiology of Learning and Memory*, 153, 26–39.

<https://doi.org/10.1016/j.nlm.2018.02.017>

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>

Bransford, J. D., Marcia, & Johnson, K. (1972). Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717–726.

Breen, K. C., & Drutte, G. (2013). Non-motor symptoms of Parkinson's disease: The patient's perspective. *Journal of Neural Transmission*, 120(4), 531–535. <https://doi.org/10.1007/s00702-012-0928-2>

Brewer, W. F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (Eds.), *Remembering Reconsidered* (1st ed., pp. 21–90). Cambridge University Press. <https://doi.org/10.1017/CBO9780511664014.004>

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, 13(2), 207–230. [https://doi.org/10.1016/0010-0285\(81\)90008-6](https://doi.org/10.1016/0010-0285(81)90008-6)

Brod, G., Lindenberger, U., Werkle-Bergner, M., & Shing, Y. L. (2015). Differences in the neural signature of remembering schema-congruent and schema-incongruent events. *NeuroImage*, 117, 358–366. <https://doi.org/10.1016/j.neuroimage.2015.05.086>

Brod, G., & Shing, Y. L. (2018). Specifying the role of the ventromedial prefrontal cortex in memory formation. *Neuropsychologia*, 111, 8–15. <https://doi.org/10.1016/j.neuropsychologia.2018.01.005>

Brod, G., Werkle-Bergner, M., & Shing, Y. L. (2013). The Influence of Prior Knowledge on Memory: A Developmental Cognitive Neuroscience Perspective. *Frontiers in Behavioral Neuroscience*, 7. <https://doi.org/10.3389/fnbeh.2013.00139>

- Brown, N. R. (2016). Transition Theory: A Minimalist Perspective on the Organization of Autobiographical Memory. *Journal of Applied Research in Memory and Cognition*, 5.
- Brown, N. R. (2021). The possible effects of the COVID-19 pandemic on the contents and organization of autobiographical memory: A Transition-Theory perspective. *Cognition*, 212, 104694.  
<https://doi.org/10.1016/j.cognition.2021.104694>
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5(1), 73–99.  
[https://doi.org/10.1016/0010-0277\(77\)90018-X](https://doi.org/10.1016/0010-0277(77)90018-X)
- Brunec, I. K., Ozubko, J. D., Ander, T., Guo, R., Moscovitch, M., & Barese, M. D. (2020). Turns during navigation act as boundaries that enhance spatial memory and expand time estimation. *Neuropsychologia*, 141, 107437. <https://doi.org/10.1016/j.neuropsychologia.2020.107437>
- Brydges, C. R., & Bielak, A. A. M. (2020). A Bayesian Analysis of Evidence in Support of the Null Hypothesis in Gerontological Psychology (or Lack Thereof). *The Journals of Gerontology: Series B*, 75(1), 58–66. <https://doi.org/10.1093/geronb/gbz033>
- Bubic, A., Von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4. <https://doi.org/10.3389/fnhum.2010.00025>
- Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, 11(2), 49–57. <https://doi.org/10.1016/j.tics.2006.11.004>
- Buonomano, D., Buzsáki, G., Davachi, L., & Nobre, A. (2023). Time for Memories. *The Journal of Neuroscience*, 43, 7565–7574. <https://doi.org/10.1523/JNEUROSCI.1430-23.2023>
- Burdett, B. R. D., Starkey, N. J., & Charlton, S. G. (2018). Characteristics of the close to home crash. *Safety Science*, 105, 222–227. <https://doi.org/10.1016/j.ssci.2018.02.020>
- Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), Article 1. <https://doi.org/10.18637/jss.v080.i01>
- Burt, C. D. B., Kemp, S., & Conway, M. A. (2003). Themes, events, and episodes in autobiographical memory. *Memory & Cognition*, 31(2), 317–325. <https://doi.org/10.3758/BF03194390>
- Castellà, J., Ueno, T., & Allen, R. J. (2024). 2020 feels slow, long, and far away: Time distortion due to

the COVID-19 pandemic. *Applied Cognitive Psychology*, 38(2), e4182.

<https://doi.org/10.1002/acp.4182>

Charlton, S. G., & Starkey, N. J. (2013). Driving on familiar roads: Automaticity and inattention blindness. *Transportation Research Part F: Traffic Psychology and Behaviour*, 19, 121–133.

<https://doi.org/10.1016/j.trf.2013.03.008>

Chaumon, M., Rioux, P.-A., Herbst, S. K., Spiousas, I., Kübel, S. L., Gallego Hiroyasu, E. M., Runyun, S. L., Micillo, L., Thanopoulos, V., Mendoza-Duran, E., Wagelmans, A., Mudumba, R., Tachmatzidou, O., Cellini, N., D'Argembeau, A., Giersch, A., Grondin, S., Gronfier, C., Igarzábal, F. A., ... van Wassenhove, V. (2022). The Blursday database as a resource to study subjective temporalities during COVID-19. *Nature Human Behaviour*, 6(11), 1587–1599.

<https://doi.org/10.1038/s41562-022-01419-2>

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

Chiew, K. S., Harris, B. B., & Adcock, R. A. (2022). Remembering Election Night 2016: Subjective but not objective metrics of autobiographical memory vary with political affiliation, affective valence, and surprise. *Journal of Experimental Psychology: General*, 151(2), 390–409.

<https://doi.org/10.1037/xge0001080>

Chow, T. E., & Rissman, J. (2017). Neurocognitive mechanisms of real-world autobiographical memory retrieval: Insights from studies using wearable camera technology. *Annals of the New York Academy of Sciences*, 1396(1), 202–221. <https://doi.org/10.1111/nyas.13353>

Chung, Y.-C., Fisher, B. E., Finley, J. M., Kim, A., Petkus, A. J., Schiehser, D. M., Jakowec, M. W., & Petzinger, G. M. (2021). Cognition and motor learning in a Parkinson's disease cohort: Importance of recall in episodic memory. *NeuroReport*, 32(14), 1153.

<https://doi.org/10.1097/WNR.0000000000001707>

Clewett, D., & Davachi, L. (2017). The ebb and flow of experience determines the temporal structure of

memory. *Current Opinion in Behavioral Sciences*, 17, 186–193.

<https://doi.org/10.1016/j.cobeha.2017.08.013>

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>

Clewett, D., Dunsmoor, J., Bachman, S. L., Phelps, E. A., & Davachi, L. (2022). Survival of the salient: Aversive learning rescues otherwise forgettable memories via neural reactivation and post-encoding hippocampal connectivity. *Neurobiology of Learning and Memory*, 187, 107572.  
<https://doi.org/10.1016/j.nlm.2021.107572>

Clewett, D., Gasser, C., & Davachi, L. (2020). Pupil-linked arousal signals track the temporal organization of events in memory. *Nature Communications*, 11(1).

<https://doi.org/10.1038/s41467-020-17851-9>

Cliff, N. (1996). *Ordinal Methods for Behavioral Data Analysis*. Psychology Press.

Cliver, K. G., Gregory, D. F., Martinez, S. A., Mitchell, W. J., Stasiak, J. E., Reisman, S. S., Helion, C., & Murty, V. P. (2024). Temporal memory for threatening events encoded in a haunted house. *Cognition & Emotion*, 1–17. <https://doi.org/10.1080/02699931.2024.2338962>

Cohen, J. (1973). Eta-Squared and Partial Eta-Squared in Fixed Factor Anova Designs. *Educational and Psychological Measurement*, 33(1), 107–112. <https://doi.org/10.1177/001316447303300111>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Corkin, S. U. (1968). Acquisition of motor skill after bilateral medial temporal lobe excision. *Neuropsychologia*, 225–265.

Coutanche, M. N., & Thompson-Schill, S. L. (2014). Fast mapping rapidly integrates information into existing memory networks. *Journal of Experimental Psychology. General*, 143(6), 2296–2303.

<https://doi.org/10.1037/xge0000020>

Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The Necessity of the Hippocampus for Statistical Learning. *Journal of Cognitive Neuroscience*, 30(5), 680–697.

[https://doi.org/10.1162/jocn\\_a\\_01228](https://doi.org/10.1162/jocn_a_01228)

- Cowan, E. T., Fain, M., O’Shea, I., Ellman, L. M., & Murty, V. P. (2021). VTA and Anterior Hippocampus Target Dissociable Neocortical Networks for Post-Novelty Enhancements. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.0316-21.2021>
- Cox, C. D., Palmer, L. C., Pham, D. T., Trieu, B. H., Gall, C. M., & Lynch, G. (2017). Experiential learning in rodents: Past experience enables rapid learning and localized encoding in hippocampus. *Learning & Memory*, 24(11), 569–579. <https://doi.org/10.1101/lm.045559.117>
- Crovitz, H. F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society*, 4(5), 517–518. <https://doi.org/10.3758/BF03334277>
- Crowe, D. A., Zarco, W., Bartolo, R., & Merchant, H. (2014). Dynamic Representation of the Temporal and Sequential Structure of Rhythmic Movements in the Primate Medial Premotor Cortex. *Journal of Neuroscience*, 34(36), 11972–11983. <https://doi.org/10.1523/JNEUROSCI.2177-14.2014>
- Curran, T. (1997). Higher-Order Associative Learning in Amnesia: Evidence from the Serial Reaction Time Task. *Journal of Cognitive Neuroscience*, 9(4), 522–533.  
<https://doi.org/10.1162/jocn.1997.9.4.522>
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction. *NeuroImage*, 9(2), 179–194. <https://doi.org/10.1006/nimg.1998.0395>
- 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>
- Davis, C. D., Jones, F. L., & Derrick, B. E. (2004). Novel Environments Enhance the Induction and Maintenance of Long-Term Potentiation in the Dentate Gyrus. *Journal of Neuroscience*, 24(29), 6497–6506. <https://doi.org/10.1523/JNEUROSCI.4970-03.2004>
- De Oliveira Alvares, L., Crestani, A. P., Cassini, L. F., Haubrich, J., Santana, F., & Quillfeldt, J. A. (2013). Reactivation enables memory updating, precision-keeping and strengthening: Exploring the possible biological roles of reconsolidation. *Neuroscience*, 244, 42–48.  
<https://doi.org/10.1016/j.neuroscience.2013.04.005>

- deBettencourt, M. T., Norman, K. A., & Turk-Browne, N. B. (2018). Forgetting from lapses of sustained attention. *Psychonomic Bulletin & Review*, 25(2), 605–611. <https://doi.org/10.3758/s13423-017-1309-5>
- Den Ouden, H. E., Kok, P., & De Lange, F. P. (2012). How Prediction Errors Shape Perception, Attention, and Motivation. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00548>
- Desrochers, T. M., Chatham, C. H., & Badre, D. (2015). The Necessity of Rostrolateral Prefrontal Cortex for Higher-Level Sequential Behavior. *Neuron*, 87(6), 1357–1368. <https://doi.org/10.1016/j.neuron.2015.08.026>
- Desrochers, T. M., Collins, A. G. E., & Badre, D. (2019). Sequential Control Underlies Robust Ramping Dynamics in the Rostrolateral Prefrontal Cortex. *The Journal of Neuroscience*, 39(8), 1471–1483. <https://doi.org/10.1523/JNEUROSCI.1060-18.2018>
- Dev, D. K., Wardell, V., Checknita, K. J., Te, A. A., Petrucci, A. S., Le, M. L., Madan, C. R., & Palombo, D. J. (2022). Negative emotion enhances memory for the sequential unfolding of a naturalistic experience. *Journal of Applied Research in Memory and Cognition*, 11(4), 510–521. <https://doi.org/10.1037/mac0000015>
- DeVito, L. M., & Eichenbaum, H. (2011). Memory for the order of events in specific sequences: Contributions of the hippocampus and medial prefrontal cortex. *Journal of Neuroscience*, 31(9), 3169–3175. <https://doi.org/10.1523/JNEUROSCI.4202-10.2011>
- DeWitt, M. R., Knight, J. B., Hicks, J. L., & Ball, B. H. (2012). The effects of prior knowledge on the encoding of episodic contextual details. *Psychonomic Bulletin & Review*, 19(2), 251–257. <https://doi.org/10.3758/s13423-011-0196-4>
- Diamond, N. B., & Levine, B. (2020). Linking Detail to Temporal Structure in Naturalistic-Event Recall. *Psychological Science*, 31(12), 1557–1572. <https://doi.org/10.1177/0956797620958651>
- Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4), 805–815. <https://doi.org/10.1037/0278-7393.32.4.805>

- Dobbins, I. G., Schnyer, D. M., Verfaellie, M., & Schacter, D. L. (2004). Cortical activity reductions during repetition priming can result from rapid response learning. *Nature*, 428(6980), 316–319.  
<https://doi.org/10.1038/nature02400>
- Doherty, A. R., Pauly-Takacs, K., Caprani, N., Gurrin, C., Moulin, C. J. A., O'Connor, N. E., & Smeaton, A. F. (2012). Experiences of Aiding Autobiographical Memory Using the SenseCam. *Human–Computer Interaction*, 27(1–2), 151–174. <https://doi.org/10.1080/07370024.2012.656050>
- Dubourg, L., Silva, A. R., Fitamen, C., Moulin, C. J. A., & Souchay, C. (2016). SenseCam: A new tool for memory rehabilitation? *Revue Neurologique*, 172(12), 735–747.  
<https://doi.org/10.1016/j.neurol.2016.03.009>
- 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., & Davachi, L. (2014). Temporal Memory Is Shaped by Encoding Stability and Intervening Item Reactivation. *Journal of Neuroscience*, 34(42). <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>
- 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>
- DuBrow, S., Sherman, B. E., Meager, M. R., & Davachi, L. (2024). Medial Temporal Lobe Damage Impairs Temporal Integration in Episodic Memory. *Journal of Cognitive Neuroscience*, 1–15.  
[https://doi.org/10.1162/jocn\\_a\\_02222](https://doi.org/10.1162/jocn_a_02222)
- Dudai, Y., Karni, A., & Born, J. (2015). The Consolidation and Transformation of Memory. *Neuron*, 88(1), 20–32. <https://doi.org/10.1016/j.neuron.2015.09.004>
- Duncan, K. D., 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>

Dunsmoor, J. E., Murty, V. P., Clewett, D., Phelps, E. A., & Davachi, L. (2022). Tag and capture: How salient experiences target and rescue nearby events in memory. *Trends in Cognitive Sciences*,

26(9), 782–795. <https://doi.org/10.1016/j.tics.2022.06.009>

Dunsmoor, J. E., Murty, V. P., Davachi, L., & Phelps, E. A. (2015). Emotional learning selectively and retroactively strengthens memories for related events. *Nature*, 520(7547), 345–348.

<https://doi.org/10.1038/nature14106>

Duszkiewicz, A. J., McNamara, C. G., Takeuchi, T., & Genzel, L. (2019). Novelty and Dopaminergic Modulation of Memory Persistence: A Tale of Two Systems. *Trends in Neurosciences*, 42(2), 102–114. <https://doi.org/10.1016/j.tins.2018.10.002>

Düzel, E., Bunzeck, N., Guitart-Masip, M., & Düzel, S. (2010). NOvelty-related Motivation of Anticipation and exploration by Dopamine (NOMAD): Implications for healthy aging. *Neuroscience & Biobehavioral Reviews*, 34(5), 660–669.

<https://doi.org/10.1016/j.neubiorev.2009.08.006>

Ecker, U. K. H., Zimmer, H. D., & Groh-Bordin, C. (2007). Color and context: An ERP study on intrinsic and extrinsic feature binding in episodic memory. *Memory & Cognition*, 35(6), 1483–1501.

<https://doi.org/10.3758/BF03193618>

Eichenbaum, H. (2014). *Time cells in the hippocampus: A new dimension for mapping memories*. 15(11), 732–744. <https://doi.org/10.1038/nrn3827>

Eichenbaum, H. (2017). Time (and space) in the hippocampus. *Current Opinion in Behavioral Sciences*, 17, 65–70. <https://doi.org/10.1016/j.cobeha.2017.06.010>

Engelkamp, J. (1998). *Memory for actions* (pp. vii, 166). Psychology Press/Taylor & Francis (UK).

Engelkamp, J., & Cohen, R. L. (1991). Current issues in memory of action events. *Psychological Research*, 53(3), 175–182. <https://doi.org/10.1007/BF00941384>

Engelkamp, J., Zimmer, H. D., Mohr, G., & Sellen, O. (1994). Memory of self-performed tasks: Self-performing during recognition. *Memory & Cognition*, 22(1), 34–39.

<https://doi.org/10.3758/BF03202759>

Ergo, K., Loof, E. D., & Verguts, T. (2020). Reward Prediction Error and Declarative Memory. *Trends in Cognitive Sciences*, 24(5), 388–397. <https://doi.org/10.1016/j.tics.2020.02.009>

Eslinger, P. J., & Damasio, A. R. (1986). Preserved motor learning in Alzheimer's disease: Implications for anatomy and behavior. *Journal of Neuroscience*, 6(10), 3006–3009.

<https://doi.org/10.1523/JNEUROSCI.06-10-03006.1986>

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

Esteban, O., Markiewicz, C. J., Goncalves, M., DuPre, E., Kent, J. D., Salo, T., Ceric, R., Pinsard, B., Blair, R. W., Poldrack, R. A., & Gorgolewski, K. J. (2022). *fMRIprep: A robust preprocessing pipeline for functional MRI* (Version 20.2.7) [Computer software]. Zenodo.

<https://doi.org/10.5281/zenodo.5898597>

Evans, A. C., Janke, A. L., Collins, D. L., & Baillet, S. (2012). Brain templates and atlases. *NeuroImage*, 62(2), 911–922. <https://doi.org/10.1016/j.neuroimage.2012.01.024>

Exner, C., Koschack, J., & Irle, E. (2002). The Differential Role of Premotor Frontal Cortex and Basal Ganglia in Motor Sequence Learning: Evidence From Focal Basal Ganglia Lesions. *Learning & Memory*, 9(6), 376–386. <https://doi.org/10.1101/lm.48402>

Ezzyat, Y., & Davachi, L. (2011). What constitutes an episode in episodic memory? *Psychological Science*, 22(2), 243–252. <https://doi.org/10.1177/0956797610393742>

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., & Davachi, L. (2021). Neural evidence for representational persistence within events. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.0073-21.2021>

Fairhall, S. L., & Caramazza, A. (2013). Brain Regions That Represent Amodal Conceptual Knowledge. *Journal of Neuroscience*, 33(25), 10552–10558. <https://doi.org/10.1523/JNEUROSCI.0051-13.2013>

Farovik, A., Dupont, L. M., & Eichenbaum, H. (2010). Distinct roles for dorsal CA3 and CA1 in memory for sequential nonspatial events. *Learning & Memory*, 17(1), 801–806. <https://doi.org/10.1101/lm.1616209>

Farzanfar, D., Spiers, H. J., Moscovitch, M., & Rosenbaum, R. S. (2023). From cognitive maps to spatial schemas. *Nature Reviews Neuroscience*, 24(2), 63–79. <https://doi.org/10.1038/s41583-022-00655-9>

Fenker, D. B., Frey, J. U., Schuetze, H., Heipertz, D., Heinze, H.-J., & Duzel, E. (2008). Novel Scenes Improve Recollection and Recall of Words. *Journal of Cognitive Neuroscience*, 20(7), 1250–1265. <https://doi.org/10.1162/jocn.2008.20086>

Finley, J. R., & Brewer, W. F. (2024). Accuracy and completeness of autobiographical memory: Evidence from a wearable camera study. *Memory*, 1–31. <https://doi.org/10.1080/09658211.2024.2377193>

Fischl, B. (2012). FreeSurfer. *NeuroImage*, 62(2), 774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>

Folville, A., Simons, J. S., D'Argembeau, A., & Bastin, C. (2022). I remember it like it was yesterday: Age-related differences in the subjective experience of remembering. *Psychonomic Bulletin & Review*, 29(4), 1223–1245. <https://doi.org/10.3758/s13423-021-02048-y>

Folville, A., Willems, S., Cheriet, N., Geurten, M., Guillemin, C., Muto, V., Requier, F., Reyt, M., Collette, F., Schmidt, C., & Bastin, C. (2023). Well-being during COVID-19-related first lockdown: Relationship with autobiographical memory and experiential diversity. *Applied Cognitive Psychology*, 37(5), 1059–1070. <https://doi.org/10.1002/acp.4104>

Fonov, V., Evans, A., McKinstry, R., Almlie, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47, S102. [https://doi.org/10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5)

- Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, 5(5), 458–462. <https://doi.org/10.1038/nn834>
- Frank, D., & Kafkas, A. (2021). Expectation-driven novelty effects in episodic memory. *Neurobiology of Learning and Memory*, 183, 107466. <https://doi.org/10.1016/j.nlm.2021.107466>
- Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences*, 17(6), 791–802. [https://doi.org/10.1016/0191-8869\(94\)90048-5](https://doi.org/10.1016/0191-8869(94)90048-5)
- Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, 25(4), 355–373. [https://doi.org/10.1016/S0149-7634\(01\)00019-7](https://doi.org/10.1016/S0149-7634(01)00019-7)
- Fuster, J. M. (2001). The Prefrontal Cortex—An Update: Time Is of the Essence. *Neuron*, 30(2), 319–333. [https://doi.org/10.1016/S0896-6273\(01\)00285-9](https://doi.org/10.1016/S0896-6273(01)00285-9)
- Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 37(7), 1297–1310. <https://doi.org/10.1016/j.neubiorev.2013.03.023>
- Gasser, C., & Davachi, L. (2023). Cross-Modal Facilitation of Episodic Memory by Sequential Action Execution. *Psychological Science*, 34(5), 581–602. <https://doi.org/10.1177/09567976231158292>
- Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia*, 53(1), 104–114. <https://doi.org/10.1016/j.neuropsychologia.2013.11.010>
- 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., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C., & Jenkinson, M. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, 80, 105–124. <https://doi.org/10.1016/j.neuroimage.2013.04.127>

- Gold, P. E. (2003). Acetylcholine modulation of neural systems involved in learning and memory. *Neurobiology of Learning and Memory*, 80(3), 194–210.  
<https://doi.org/10.1016/j.nlm.2003.07.003>
- González-Cutre, D., Romero-Elías, M., Jiménez-Loaisa, A., Beltrán-Carrillo, V. J., & Hagger, M. S. (2020). Testing the need for novelty as a candidate need in basic psychological needs theory. *Motivation and Emotion*, 44(2), 295–314. <https://doi.org/10.1007/s11031-019-09812-7>
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). NiPy: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*, 5. <https://doi.org/10.3389/fninf.2011.00013>
- Graesser, A. C., & Nakamura, G. V. (1982). The Impact of a Schema on Comprehension and Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 16, pp. 59–109). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60547-2](https://doi.org/10.1016/S0079-7421(08)60547-2)
- Graybiel, A. M., & Grafton, S. T. (2015). The Striatum: Where Skills and Habits Meet. *Cold Spring Harbor Perspectives in Biology*, 7(8), a021691. <https://doi.org/10.1101/cshperspect.a021691>
- Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, 94, 149–165.  
<https://doi.org/10.1016/j.jml.2016.11.001>
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1), 63–72. <https://doi.org/10.1016/j.neuroimage.2009.06.060>
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10), 1409–1422. [https://doi.org/10.1016/S0042-6989\(01\)00073-6](https://doi.org/10.1016/S0042-6989(01)00073-6)
- Grön, G., Kirstein, M., Thielscher, A., Riepe, M. W., & Spitzer, M. (2005). Cholinergic enhancement of episodic memory in healthy young adults. *Psychopharmacology*, 182(1), 170–179.  
<https://doi.org/10.1007/s00213-005-0043-2>
- Gruber, M. J., Ritchey, M., Wang, S.-F., Doss, M. K., & Ranganath, C. (2016). Post-learning

- Hippocampal Dynamics Promote Preferential Retention of Rewarding Events. *Neuron*, 89(5), 1110–1120. <https://doi.org/10.1016/j.neuron.2016.01.017>
- Guo, D., Chen, H., Wang, L., & Yang, J. (2023). Effects of prior knowledge on brain activation and functional connectivity during memory retrieval. *Scientific Reports*, 13(1), 13650. <https://doi.org/10.1038/s41598-023-40966-0>
- Guo, D., & Yang, J. (2020). Interplay of the long axis of the hippocampus and ventromedial prefrontal cortex in schema-related memory retrieval. *Hippocampus*, 30(3), 263–277. <https://doi.org/10.1002/hipo.23154>
- Guo, D., & Yang, J. (2023). Reactivation of schema representation in lateral occipital cortex supports successful memory encoding. *Cerebral Cortex*, 33(10), 5968–5980. <https://doi.org/10.1093/cercor/bhac475>
- Habib, R., McIntosh, A. R., Wheeler, M. A., & Tulving, E. (2003). Memory encoding and hippocampally-based novelty/familiarity discrimination networks. *Neuropsychologia*, 41(3), 271–279. [https://doi.org/10.1016/S0028-3932\(02\)00160-4](https://doi.org/10.1016/S0028-3932(02)00160-4)
- Harms, I. M. (2023). Distracted by familiarity: Implications of ‘autopilot’ as a default cognitive mode. *Transportation Research Part F: Traffic Psychology and Behaviour*, 99, 274–288. <https://doi.org/10.1016/j.trf.2023.10.021>
- Hasselmo, M. E. (2006). The role of acetylcholine in learning and memory. *Current Opinion in Neurobiology*, 16(6), 710–715. <https://doi.org/10.1016/j.conb.2006.09.002>
- Hasselmo, M. E., & Stern, C. E. (2014). Theta rhythm and the encoding and retrieval of space and time. 85, 656–666. <https://doi.org/10.1016/j.neuroimage.2013.06.022>
- Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d'. *Behavior Research Methods, Instruments, & Computers*, 27(1), 46–51. <https://doi.org/10.3758/BF03203619>
- Healey, M. K., Long, N. M., & Kahana, M. J. (2019). Contiguity in episodic memory. *Psychonomic Bulletin & Review*, 26(3), 699–720. <https://doi.org/10.3758/s13423-018-1537-3>

- Heller, A. S., Shi, T. C., Ezie, C. E. C., Reneau, T. R., Baez, L. M., Gibbons, C. J., & Hartley, C. A. (2020). Association between real-world experiential diversity and positive affect relates to hippocampal–striatal functional connectivity. *Nature Neuroscience*, 23(7), 800–804.  
<https://doi.org/10.1038/s41593-020-0636-4>
- Henderson, L. M., & James, E. (2018). Consolidating new words from repetitive versus multiple stories: Prior knowledge matters. *Journal of Experimental Child Psychology*, 166, 465–484.  
<https://doi.org/10.1016/j.jecp.2017.09.017>
- Hendriks, M. H. A., Daniels, N., Pegado, F., & Op de Beeck, H. P. (2017). The Effect of Spatial Smoothing on Representational Similarity in a Simple Motor Paradigm. *Frontiers in Neurology*, 8. <https://doi.org/10.3389/fneur.2017.00222>
- Henin, S., Turk-Browne, N. B., Friedman, D., Liu, A., Dugan, P., Flinker, A., Doyle, W., Devinsky, O., & Melloni, L. (2021). Learning hierarchical sequence representations across human cortex and hippocampus. *Science Advances*, 7(8), eabc4530. <https://doi.org/10.1126/sciadv.abc4530>
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *The British Journal of Clinical Psychology*, 44(Pt 2), 227–239. <https://doi.org/10.1348/014466505X29657>
- Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*, 20(11), 1315–1326. <https://doi.org/10.1002/hipo.20857>
- Hermans, E. J., Kanen, J. W., Tambini, A., Fernández, G., Davachi, L., & Phelps, E. A. (2017). Persistence of Amygdala–Hippocampal Connectivity and Multi-Voxel Correlation Structures During Awake Rest After Fear Learning Predicts Long-Term Expression of Fear. *Cerebral Cortex*, 27(5), 3028–3041. <https://doi.org/10.1093/cercor/bhw145>
- 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>

- Hirst, W., & Phelps, E. A. (2016). Flashbulb Memories. *Current Directions in Psychological Science*, 25(1), 36–41. <https://doi.org/10.1177/0963721415622487>
- Hirst, W., Phelps, E. A., Meksin, R., Vaidya, C. J., Johnson, M. K., Mitchell, K. J., Buckner, R. L., Budson, A. E., Gabrieli, J. D. E., Lustig, C., Mather, M., Ochsner, K. N., Schacter, D., Simons, J. S., Lyle, K. B., Cuc, A. F., & Olsson, A. (2015). A ten-year follow-up of a study of memory for the attack of September 11, 2001: Flashbulb memories and memories for flashbulb events. *Journal of Experimental Psychology. General*, 144(3), 604–623.  
<https://doi.org/10.1037/xge0000055>
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8(11), 494–500. <https://doi.org/10.1016/j.tics.2004.08.007>
- Hope, R. (2022). Rmisc: A collection of functions useful for data analysis and utility operations. *R Package Version 1.5.1*. <https://doi.org/10.5281/zenodo.683128>
- Hope, R. M. (2022). *Rmisc: Ryan Miscellaneous* (Version 1.5.1) [Computer software]. <https://cran.r-project.org/web/packages/Rmisc/index.html>
- Hopkins, R. O., Waldram, K., & Kesner, R. P. (2004). Sequences assessed by declarative and procedural tests of memory in amnesic patients with hippocampal damage. *Neuropsychologia*, 42(14), 1877–1886. <https://doi.org/10.1016/j.neuropsychologia.2004.05.008>
- Horner, A. J., Bisby, J. A., Bush, D., Lin, W.-J., & Burgess, N. (2015). Evidence for holistic episodic recollection via hippocampal pattern completion. *Nature Communications*, 6(1), Article 1.  
<https://doi.org/10.1038/ncomms8462>
- 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>
- Horstmann, G. (2002). Evidence for attentional capture by a surprising color singleton in visual search. *Psychological Science*, 13(6), 499–505. <https://doi.org/10.1111/1467-9280.00488>
- 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), 1211–1230.

<https://doi.org/10.1037/a0033621>

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, 268–299.

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>

Hu, X., Cheng, L. Y., Chiu, M. H., & Paller, K. A. (2020). Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological Bulletin*, 146(3), 218–244.

<https://doi.org/10.1037/bul0000223>

Huang, J., Velarde, I., Ma, W. J., & Baldassano, C. (2023). Schema-based predictive eye movements support sequential memory encoding. *eLife*, 12, e82599. <https://doi.org/10.7554/eLife.82599>

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review*, 2(1), 105–112. <https://doi.org/10.3758/BF03214414>

Irish, M., Cunningham, C. J., Walsh, J. B., Coakley, D., Lawlor, B. A., Robertson, I. H., & Coen, R. F. (2006). Investigating the Enhancing Effect of Music on Autobiographical Memory in Mild Alzheimer's Disease. *Dementia and Geriatric Cognitive Disorders*, 22(1), 108–120.

<https://doi.org/10.1159/000093487>

Jääskeläinen, I. P., Sams, M., Glerean, E., & Ahveninen, J. (2021). Movies and narratives as naturalistic stimuli in neuroimaging. *NeuroImage*, 224, 117445.

<https://doi.org/10.1016/j.neuroimage.2020.117445>

Jacobacci, F., Armony, J. L., Yeffal, A., Lerner, G., Amaro, E., Jovicich, J., Doyon, J., & Della-

Maggiore, V. (2020). Rapid hippocampal plasticity supports motor sequence learning.

*Proceedings of the National Academy of Sciences*, 117(38), 23898–23903.

<https://doi.org/10.1073/pnas.2009576117>

Janssen, S. M. J., Chessa, A. G., & Murre, J. M. J. (2006). Memory for time: How people date events.

*Memory & Cognition*, 34(1), 138–147. <https://doi.org/10.3758/BF03193393>

Jardine, K. H., Huff, A. E., Wideman, C. E., McGraw, S. D., & Winters, B. D. (2022). The evidence for and against reactivation-induced memory updating in humans and nonhuman animals.

*Neuroscience & Biobehavioral Reviews*, 136, 104598.

<https://doi.org/10.1016/j.neubiorev.2022.104598>

Jayakumar, M., Balusu, C., & Aly, M. (2023). Attentional fluctuations and the temporal organization of memory. *Cognition*, 235, 105408. <https://doi.org/10.1016/j.cognition.2023.105408>

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>

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

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

Jeunehomme, O., Heinen, R., Stawarczyk, D., Axmacher, N., & D'Argembeau, A. (2022). Representational dynamics of memories for real-life events. *iScience*, 25(11), 105391.

<https://doi.org/10.1016/j.isci.2022.105391>

Kafkas, A., & Montaldi, D. (2015). Striatal and midbrain connectivity with the hippocampus selectively boosts memory for contextual novelty. *Hippocampus*, 25(11), 1262–1273.

<https://doi.org/10.1002/hipo.22434>

Kafkas, A., & Montaldi, D. (2018). How do memory systems detect and respond to novelty? *Neuroscience Letters*, 680, 60–68. <https://doi.org/10.1016/j.neulet.2018.01.053>

Kahana, M. J., Howard, M. W., & Polyn, S. M. (2008). Associative Retrieval Processes in Episodic Memory. In *Learning and Memory: A Comprehensive Reference* (pp. 1–24). Elsevier.

<https://doi.org/10.1016/B978-012370509-9.00185-6>

Kalm, K., Davis, M. H., & Norris, D. (2013). Individual sequence representations in the medial temporal lobe. *Journal of Cognitive Neuroscience*, 25(7), 1111–1121. <https://doi.org/10.1162/jocn>

Kassambara, A. (2021). rstatix: Pipe-friendly framework for basic statistical tests. *R Package Version 0.7.0*. <https://rpkgs.datanovia.com/rstatix/>

Kesner, R. P. (2013). Role of the hippocampus in mediating interference as measured by pattern separation processes. *Behavioural Processes*, 93, 148–154.

<https://doi.org/10.1016/j.beproc.2012.09.018>

Kesner, R. P., Gilbert, P. E., & Barua, L. A. (2002). The role of the hippocampus in memory for the temporal order of a sequence of odors. *Behavioral Neuroscience*, 116(2), 286–290.

<https://doi.org/10.1037/0735-7044.116.2.286>

Kinder, K. T., & Buss, A. T. (2021). The effect of motor engagement on memory: Testing a motor-induced encoding account. *Memory & Cognition*, 49(3), 586–599.

<https://doi.org/10.3758/s13421-020-01113-6>

King, B. R., Dolfen, N., Gann, M. A., Renard, Z., Swinnen, S. P., & Albouy, G. (2019). Schema and Motor-Memory Consolidation. *Psychological Science*, 30(7), 963–978.

<https://doi.org/10.1177/0956797619847164>

King, B. R., Gann, M. A., Mantini, D., Doyon, J., & Albouy, G. (2022). Persistence of hippocampal and striatal multivoxel patterns during awake rest after motor sequence learning. *iScience*, 25(12), 105498. <https://doi.org/10.1016/j.isci.2022.105498>

Kleider, H. M., Pezdek, K., Goldinger, S. D., & Kirk, A. (2008). Schema-driven source misattribution errors: Remembering the expected from a witnessed event. *Applied Cognitive Psychology*, 22(1), 1–20. <https://doi.org/10.1002/acp.1361>

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

Kole, J. A., & Healy, A. F. (2007). Using prior knowledge to minimize interference when learning large amounts of information. *Memory & Cognition*, 35(1), 124–137.

<https://doi.org/10.3758/BF03195949>

Kormi-Nouri, R., Nilsson, L., & Ohta, N. (2005). The novelty effect: Support for the Novelty-Encoding Hypothesis. *Scandinavian Journal of Psychology*, 46(2), 133–143. <https://doi.org/10.1111/j.1467-9450.2005.00443.x>

Kornysheva, K., & Diedrichsen, J. (2014). Human premotor areas parse sequences into their spatial and temporal features. *eLife*, 3, e03043. Scopus. <https://doi.org/10.7554/eLife.03043>

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2, 4.

<https://doi.org/10.3389/neuro.06.004.2008>

Kuhl, B. A., Bainbridge, W. A., & Chun, M. M. (2012). Neural Reactivation Reveals Mechanisms for Updating Memory. *Journal of Neuroscience*, 32(10), 3453–3461.

<https://doi.org/10.1523/JNEUROSCI.5846-11.2012>

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. 12(2), 72–79.

<https://doi.org/10.1016/j.tics.2007.11.004>

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82, 1–26.

<https://doi.org/10.18637/jss.v082.i13>

Kyle, C. T., Smuda, D. N., Hassan, A. S., & Ekstrom, A. D. (2015). Roles of human hippocampal subfields in retrieval of spatial and temporal context. 278, 549–558.

<https://doi.org/10.1016/j.bbr.2014.10.034>

La Corte, V., Sperduti, M., Abichou, K., & Piolino, P. (2019). Episodic Memory Assessment and Remediation in Normal and Pathological Aging Using Virtual Reality: A Mini Review. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00173>

Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied*

*Mathematics Series B Numerical Analysis*, 1(1), 76–85. <https://doi.org/10.1137/0701007>

Larson, R., & Csikszentmihalyi, M. (2014). The Experience Sampling Method. In M. Csikszentmihalyi (Ed.), *Flow and the Foundations of Positive Psychology* (pp. 21–34). Springer Netherlands. [https://doi.org/10.1007/978-94-017-9088-8\\_2](https://doi.org/10.1007/978-94-017-9088-8_2)

Lee, C. S., Aly, M., & Baldassano, C. (2021). Anticipation of temporally structured events in the brain. *eLife*, 10, e64972. <https://doi.org/10.7554/eLife.64972>

Lee, S., Koffer, R. E., Sprague, B. N., Charles, S. T., Ram, N., & Almeida, D. M. (2016). Activity Diversity and Its Associations With Psychological Well-Being Across Adulthood. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, gbw118. <https://doi.org/10.1093/geronb/gbw118>

Lehn, H., Steffenach, H.-A., Strien, N. M. van, Veltman, D. J., Witter, M. P., & Håberg, A. K. (2009). A Specific Role of the Human Hippocampus in Recall of Temporal Sequences. *Journal of Neuroscience*, 29(11), 3475–3484. <https://doi.org/10.1523/JNEUROSCI.5370-08.2009>

Lenth, R., Love, J., & Herve, M. (2018). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (Version 1.1) [Computer software]. <https://CRAN.R-project.org/package=emmeans>

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8).

Levine, B., Svoboda, E., Hay, J. F., Winocur, G., & Moscovitch, M. (2002). Aging and autobiographical memory: Dissociating episodic from semantic retrieval. *Psychology and Aging*, 17(4), 677–689. <https://doi.org/10.1037/0882-7974.17.4.677>

Lew, A. R., & Howe, M. L. (2017). Out of place, out of mind: Schema-driven false memory effects for object-location bindings. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(3), 404–421. <https://doi.org/10.1037/xlm0000317>

Lewandowsky, S., & Murdock Jr., B. B. (1989). Memory for serial order. *Psychological Review*, 96(1), 25–57. <https://doi.org/10.1037/0033-295X.96.1.25>

Li, S., Cullen, W. K., Anwyl, R., & Rowan, M. J. (2003). Dopamine-dependent facilitation of LTP

- induction in hippocampal CA1 by exposure to spatial novelty. *Nature Neuroscience*, 6(5), 526–531. <https://doi.org/10.1038/nn1049>
- Lisman, J. E., & Grace, A. A. (2005). The Hippocampal-VTA Loop: Controlling the Entry of Information into Long-Term Memory. *Neuron*, 46(5), 703–713. <https://doi.org/10.1016/j.neuron.2005.05.002>
- Liu, C., Ye, Z., Chen, C., Axmacher, N., & Xue, G. (2021). Hippocampal Representations of Event Structure and Temporal Context during Episodic Temporal Order Memory. *Cerebral Cortex*, bhab304. <https://doi.org/10.1093/cercor/bhab304>
- Liu, Z. X., Grady, C., & Moscovitch, M. (2017). Effects of Prior-Knowledge on Brain Activation and Connectivity During Associative Memory Encoding. *Cerebral Cortex*, 27(3), 1991–2009. <https://doi.org/10.1093/cercor/bhw047>
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4(1), 19–31. <https://doi.org/10.1037/0278-7393.4.1.19>
- Logan, G. D., & Cox, G. E. (2021). Serial memory: Putting chains and position codes in context. *Psychological Review*, 128(6), 1197–1205. <https://doi.org/10.1037/rev0000327>
- Lorentz, A., Ruitenberg, M. F. L., & Schomaker, J. (2023). Novelty-induced memory boosts in humans: The when and how. *Heliyon*, 9(3), e14410. <https://doi.org/10.1016/j.heliyon.2023.e14410>
- MacDonald, C. J., Carrow, S., Place, R., & Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *Journal of Neuroscience*, 33(36), 14607–14616. <https://doi.org/10.1523/JNEUROSCI.1537-13.2013>
- Mack, M. L., Love, B. C., & Preston, A. R. (2016). Dynamic updating of hippocampal object representations reflects new conceptual knowledge. *Proceedings of the National Academy of Sciences*, 113(46), 13203–13208. <https://doi.org/10.1073/pnas.1614048113>
- Mack, M. L., Preston, A. R., & Love, B. C. (2020). Ventromedial prefrontal cortex compression during concept learning. *Nature Communications*, 11(1), 1–11. <https://doi.org/10.1038/s41467-019-13930-8>

Madan, C., & Singhal, A. (2012). Using actions to enhance memory: Effects of enactment, gestures, and exercise on human memory. *Frontiers in Psychology*, 3.

<https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00507>

Maguire, E. A., Frith, C. D., & Morris, R. G. M. (1999). The functional neuroanatomy of comprehension and memory: The importance of prior knowledge. *Brain*, 122(10), 1839–1850.

<https://doi.org/10.1093/brain/122.10.1839>

Makowski, D. (2018). The psycho Package: An Efficient and Publishing-Oriented Workflow for Psychological Science. *Journal of Open Source Software*, 3(22), 470.

<https://doi.org/10.21105/joss.00470>

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences*, 92(18), 8135–8139. <https://doi.org/10.1073/pnas.92.18.8135>

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>

Martin, C. B., Hong, B., Newsome, R. N., Savel, K., Meade, M. E., Xia, A., Honey, C. J., & Barense, M. D. (2022). A smartphone intervention that enhances real-world memory and promotes differentiation of hippocampal activity in older adults. *Proceedings of the National Academy of Sciences*, 119(51), e2214285119. <https://doi.org/10.1073/pnas.2214285119>

Masis-Obando, R., Norman, K. A., & Baldassano, C. (2022). Schema representations in distinct brain networks support narrative memory during encoding and retrieval. *eLife*, 11, e70445.

<https://doi.org/10.7554/eLife.70445>

- Mayes, A. R., Isaac, C. L., Holdstock, J. S., Hunkin, N. M., Montaldi, D., Downes, J. J., MacDonald, C., Cezayirli, E., & Roberts, J. N. (2001). Memory for single items, word pairs, and temporal order of different kinds in a patient with selective hippocampal lesions. *Cognitive Neuropsychology*, 18(2), 97–123. <https://doi.org/10.1080/02643290125897>
- Mayka, M. A., Corcos, D. M., Leurgans, S. E., & Vaillancourt, D. E. (2006). Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: A meta-analysis. *NeuroImage*, 31(4), 1453–1474. <https://doi.org/10.1016/j.neuroimage.2006.02.004>
- McCabe, J. A. (2015). Location, Location, Location! Demonstrating the Mnemonic Benefit of the Method of Loci. *Teaching of Psychology*, 42(2), 169–173. <https://doi.org/10.1177/0098628315573143>
- McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology. General*, 142(4), 1190–1210. <https://doi.org/10.1037/a0033812>
- McElhinney, M., & Annett, J. M. (1996). Pattern of Efficacy of a Musical Mnemonic on Recall of Familiar Words over Several Presentations. *Perceptual and Motor Skills*, 82(2), 395–400. <https://doi.org/10.2466/pms.1996.82.2.395>
- McGaugh, J. L. (2018). Emotional arousal regulation of memory consolidation. *Current Opinion in Behavioral Sciences*, 19, 55–60. <https://doi.org/10.1016/j.cobeha.2017.10.003>
- Melo, B., Winocur, G., & Moscovitch, M. (1999). False recall and false recognition: An examination of the effects of selective and combined lesions to the medial temporal lobe/diencephalon and frontal lobe structures. *Cognitive Neuropsychology*, 16(3–5), 343–359. Scopus. <https://doi.org/10.1080/026432999380825>
- Michelmann, S., Kumar, M., Norman, K. A., & Toneva, M. (2023). Large language models can segment narrative events similarly to humans. *arXiv*. <https://doi.org/10.48550/arXiv.2301.10297>
- Milner, B., Corkin, S., & Teuber, H.-L. (1968). Further analysis of the hippocampal amnesic syndrome: 14-year follow-up study of H.M. *Neuropsychologia*, 6(3), 215–234. [https://doi.org/10.1016/0028-3932\(68\)90021-3](https://doi.org/10.1016/0028-3932(68)90021-3)

- Moncada, D., Ballarini, F., & Viola, H. (2015). Behavioral Tagging: A Translation of the Synaptic Tagging and Capture Hypothesis. *Neural Plasticity*, 2015(1), 650780.  
<https://doi.org/10.1155/2015/650780>
- Moncada, D., & Viola, H. (2007). Induction of long-term memory by exposure to novelty requires protein synthesis: Evidence for a behavioral tagging. *The Journal of Neuroscience*, 27(28), 7476–7481.  
<https://doi.org/10.1523/JNEUROSCI.1083-07.2007>
- Morris, R. G. M. (2006). Elements of a neurobiological theory of hippocampal function: The role of synaptic plasticity, synaptic tagging and schemas. *European Journal of Neuroscience*, 23(11), 2829–2846. <https://doi.org/10.1111/j.1460-9568.2006.04888.x>
- Morton, N. W., Zippi, E. L., & Preston, A. R. (2023). Memory reactivation and suppression modulate integration of the semantic features of related memories in hippocampus. *Cerebral Cortex*, 33(14), 9020–9037. <https://doi.org/10.1093/cercor/bhad179>
- Mosha, N., & Robertson, E. M. (2016). Unstable Memories Create a High-Level Representation that Enables Learning Transfer. *Current Biology*, 26(1), 100–105.  
<https://doi.org/10.1016/j.cub.2015.11.035>
- Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage*, 59(3), 2636–2643. <https://doi.org/10.1016/j.neuroimage.2011.08.076>
- Mutanen, T. P., Bracco, M., & Robertson, E. M. (2020). A Common Task Structure Links Together the Fate of Different Types of Memories. *Current Biology*, 30(11), 2139-2145.e5. Scopus.  
<https://doi.org/10.1016/j.cub.2020.03.043>
- Nadel, L., & Moscovitch, M. (2000). Memory consolidation, retrograde amnesia, and the temporal lobe. *Current Opinion in Neurobiology*, 7. 223–250.
- Nader, K. (2003). Memory traces unbound. *Trends in Neurosciences*, 26(2), 65–72.  
[https://doi.org/10.1016/S0166-2236\(02\)00042-5](https://doi.org/10.1016/S0166-2236(02)00042-5)
- Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., & Sederberg, P. B. (2015). Human hippocampus

- represents space and time during retrieval of real-world memories. *Proceedings of the National Academy of Sciences*, 112(35), 11078–11083. <https://doi.org/10.1073/pnas.1507104112>
- Noroozian, M., Kormi-Nouri, R., Nyberg, L., & Persson, J. (2022). Hippocampal and motor regions contribute to memory benefits after enacted encoding: Cross-sectional and longitudinal evidence. *Cerebral Cortex*. <https://doi.org/10.1093/cercor/bhac262>
- OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., ... Zoph, B. (2024). *GPT-4 Technical Report* (No. arXiv:2303.08774). arXiv. <https://doi.org/10.48550/arXiv.2303.08774>
- Overschelde, J. P. V., & Healy, A. F. (2001). Learning of Nondomain Facts in High- and Low-Knowledge Domains. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(5), 1160–1171.
- Pardilla-Delgado, E., & Payne, J. D. (2017). The Deese-Roediger-McDermott (DRM) Task: A Simple Cognitive Paradigm to Investigate False Memories in the Laboratory. *Journal of Visualized Experiments : JoVE*, 119, 54793. <https://doi.org/10.3791/54793>
- Patil, A., Murty, V. P., Dunsmoor, J. E., Phelps, E. A., & Davachi, L. (2017). Reward retroactively enhances memory consolidation for related items. *Learning & Memory*, 24(1), 65–69. <https://doi.org/10.1101/lm.042978.116>
- Paz, R., Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., & Fried, I. (2010). A neural substrate in the human hippocampus for linking successive events. *Proceedings of the National Academy of Sciences*, 107(13), 6046–6051. <https://doi.org/10.1073/pnas.0910834107>
- Pearson, E., Graff, J., Bai, E., Jakubowski, K., & Belfi, A. M. (2023). Differences in autobiographical memories reported using text and voice during everyday life. *Memory*, 31(3), 393–405. <https://doi.org/10.1080/09658211.2022.2162084>
- Pettijohn, K. A., Thompson, A. N., Tamplin, A. K., Krawietz, S. A., & Radvansky, G. A. (2016). Event boundaries and memory improvement. *Cognition*, 148, 136–144.

<https://doi.org/10.1016/j.cognition.2015.12.013>

Plancher, G., Barra, J., Orriols, E., & Piolino, P. (2013). The influence of action on episodic memory: A virtual reality study. *Quarterly Journal of Experimental Psychology* (2006), 66(5), 895–909.

<https://doi.org/10.1080/17470218.2012.722657>

Polyn, S. M., & Kahana, M. J. (2008). Memory search and the neural representation of context. *Trends in Cognitive Sciences*, 12(1), 24–30. <https://doi.org/10.1016/j.tics.2007.10.010>

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>

Pooja, R., Ghosh, P., & Sreekumar, V. (2024). Towards an ecologically valid naturalistic cognitive neuroscience of memory and event cognition. *Neuropsychologia*, 203, 108970.

<https://doi.org/10.1016/j.neuropsychologia.2024.108970>

Poppenk, J., Köhler, S., & Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 36(5), 1321–1330. <https://doi.org/10.1037/a0019900>

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, 84, 320–341. <https://doi.org/10.1016/j.neuroimage.2013.08.048>

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, 23(17), R764–R773. <https://doi.org/10.1016/j.cub.2013.05.041>

Quent, J. A., & Henson, R. N. (2022). Novel immersive virtual reality experiences do not produce retroactive memory benefits for unrelated material. *Quarterly Journal of Experimental Psychology* (2006), 75(12), 2197–2210. <https://doi.org/10.1177/17470218221082491>

Quent, J. A., Henson, R. N., & Greve, A. (2021). A predictive account of how novelty influences declarative memory. *Neurobiology of Learning and Memory*, 179, 107382.

<https://doi.org/10.1016/j.nlm.2021.107382>

R Core Team. (2021). *R: A language and environment for statistical computing*. <https://www.R-project.org/>

Rainey, D. W., & Larsen, J. D. (2002). The Effect of Familiar Melodies on Initial Learning and Long-term Memory for Unconnected Text. *Music Perception*, 20(2), 173–186.  
<https://doi.org/10.1525/mp.2002.20.2.173>

Ramirez Butavand, D., Hirsch, I., Tomaiuolo, M., Moncada, D., Viola, H., & Ballarini, F. (2020). Novelty Improves the Formation and Persistence of Memory in a Naturalistic School Scenario. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00048>

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>

Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience*, 4(3), 193–202. <https://doi.org/10.1038/nrn1052>

Rasmussen, A. S., Johannessen, K. B., & Berntsen, D. (2014). Ways of sampling voluntary and involuntary autobiographical memories in daily life. *Consciousness and Cognition*, 30, 156–168.  
<https://doi.org/10.1016/j.concog.2014.09.008>

Rawson, K. A., & Van Overschelde, J. P. (2008). How does knowledge promote memory? The distinctiveness theory of skilled memory. *Journal of Memory and Language*, 58(3), 646–668.  
<https://doi.org/10.1016/j.jml.2007.08.004>

Raykov, P. P., Keidel, J. L., Oakhill, J., & Bird, C. M. (2020). The brain regions supporting schema-related processing of people's identities. *Cognitive Neuropsychology*, 37(1–2), 8–24.  
<https://doi.org/10.1080/02643294.2019.1685958>

Raykov, P. P., Keidel, J. L., Oakhill, J., & Bird, C. M. (2021). Activation of Person Knowledge in Medial Prefrontal Cortex during the Encoding of New Lifelike Events. *Cerebral Cortex*, 31(7), 3494–3505. <https://doi.org/10.1093/cercor/bhab027>

Reagh, Z. M., & Ranganath, C. (2023). Flexible reuse of cortico-hippocampal representations during encoding and recall of naturalistic events. *Nature Communications*, 14(1), 1279.

<https://doi.org/10.1038/s41467-023-36805-5>

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness.

*Learning & Memory*, 1(4), 217–229. <https://doi.org/10.1101/lm.1.4.217>

Reder, L. M., Liu, X. L., Keinath, A., & Popov, V. (2016). Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning. *Psychonomic Bulletin & Review*, 23(1), 271–277. <https://doi.org/10.3758/s13423-015-0889-1>

Reder, L. M., Victoria, L. W., Manelis, A., Oates, J. M., Dutcher, J. M., Bates, J. T., Cook, S., Aizenstein, H. J., Quinlan, J., & Gyulai, F. (2013). Why It's Easier to Remember Seeing a Face We Already Know Than One We Don't: Preexisting Memory Representations Facilitate Memory Formation. *Psychological Science*, 24(3), 363–372.

Redondo, R. L., & Morris, R. G. M. (2011). Making memories last: The synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience*, 12(1), 17–30. <https://doi.org/10.1038/nrn2963>

Reggente, N., Essoe, J. K., Baek, H. Y., & Rissman, J. (2020). The Method of Loci in Virtual Reality: Explicit Binding of Objects to Spatial Contexts Enhances Subsequent Memory Recall. *Journal of Cognitive Enhancement*, 4(1). <https://doi.org/10.1007/s41465-019-00141-8>

Reichardt, R., Polner, B., & Simor, P. (2020). Novelty Manipulations, Memory Performance, and Predictive Coding: The Role of Unexpectedness. *Frontiers in Human Neuroscience*, 14. <https://doi.org/10.3389/fnhum.2020.00152>

Reisenzein, R., Horstmann, G., & Schützwohl, A. (2019). The Cognitive-Evolutionary Model of Surprise: A Review of the Evidence. *Topics in Cognitive Science*, 11(1), 50–74. <https://doi.org/10.1111/tops.12292>

Ren, Y., Leslie, G., & Brown, T. (2024). Visual sequence encoding is modulated by music schematic structure and familiarity. *PLOS ONE*, 19(8), e0306271. <https://doi.org/10.1371/journal.pone.0306271>

Reuter, M., Rosas, H. D., & Fischl, B. (2010). Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, 53(4), 1181–1196. <https://doi.org/10.1016/j.neuroimage.2010.07.020>

- Ritchey, M., Libby, L. A., & Ranganath, C. (2015). Cortico-hippocampal systems involved in memory and cognition: The PMAT framework. In *Progress in Brain Research* (Vol. 219, pp. 45–64). Elsevier B.V. <https://doi.org/10.1016/bs.pbr.2015.04.001>
- 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>
- Robertson, E. M. (2022). Memory leaks: Information shared across memory systems. *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2022.03.010>
- Robinson, J. A. (1976). Sampling autobiographical memory. *Cognitive Psychology*, 8(4), 578–595.  
[https://doi.org/10.1016/0010-0285\(76\)90020-7](https://doi.org/10.1016/0010-0285(76)90020-7)
- Roediger, H. L., & McDermott, K. B. (1995). Creating False Memories: Remembering Words Not Presented in Lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4).
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237.  
<https://doi.org/10.3758/PBR.16.2.225>
- Rouhani, N., & Niv, Y. (2021). Signed and unsigned reward prediction errors dynamically enhance learning and memory. *eLife*, 10, e61077. <https://doi.org/10.7554/eLife.61077>
- Rouhani, N., Norman, K. A., & Niv, Y. (2018). Dissociable effects of surprising rewards on learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(9), 1430–1443. <https://doi.org/10.1037/xlm0000518>
- Rouhani, N., Stanley, D., COVID-Dynamic Team, & Adolphs, R. (2023). Collective events and individual affect shape autobiographical memory. *Proceedings of the National Academy of Sciences*, 120(29), e2221919120. <https://doi.org/10.1073/pnas.2221919120>
- Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., Eisenberg, M. L., & Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*,

129(2), 241–255. <https://doi.org/10.1016/j.cognition.2013.07.002>

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., Eickhoff, S.

B., Hakonarson, H., Gur, R. C., Gur, R. E., & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, 64, 240–256.

<https://doi.org/10.1016/j.neuroimage.2012.08.052>

Sayres, R., & Grill-Spector, K. (2008). Relating Retinotopic and Object-Selective Responses in Human Lateral Occipital Cortex. *Journal of Neurophysiology*, 100(1), 249–267.

<https://doi.org/10.1152/jn.01383.2007>

Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cerebral Cortex*, 28(9), 3095–3114.

<https://doi.org/10.1093/cercor/bhx179>

Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The Necessity of the Medial Temporal Lobe for Statistical Learning. *Journal of Cognitive Neuroscience*, 26(8), 1736–1747. [https://doi.org/10.1162/jocn\\_a\\_00578](https://doi.org/10.1162/jocn_a_00578)

Schapiro, A. C., McDevitt, E. A., Rogers, T. T., Mednick, S. C., & Norman, K. A. (2018). Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. *Nature Communications*, 9(1), Article 1. <https://doi.org/10.1038/s41467-018-06213-1>

Schapiro, A. C., Reid, A. G., Morgan, A., Manoach, D. S., Verfaellie, M., & Stickgold, R. (2019). The hippocampus is necessary for the consolidation of a task that does not require the hippocampus for initial learning. *Hippocampus*, 29(11), 1091–1100. <https://doi.org/10.1002/hipo.23101>

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society*,

372(1711). <https://doi.org/10.1098/rstb.2016.0049>

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An fMRI Study of the Role of the Medial Temporal Lobe in Implicit and Explicit Sequence Learning. *Neuron*, 37(6), 1013–1025.  
[https://doi.org/10.1016/S0896-6273\(03\)00123-5](https://doi.org/10.1016/S0896-6273(03)00123-5)

Schlichting, M. L., & Preston, A. R. (2014). Memory reactivation during rest supports upcoming learning of related content. *Proceedings of the National Academy of Sciences*, 111(44), 15845–15850.  
<https://doi.org/10.1073/pnas.1404396111>

Schlichting, M. L., & Preston, A. R. (2016). Hippocampal–medial prefrontal circuit supports memory updating during learning and post-encoding rest. *Neurobiology of Learning and Memory*, 134, 91–106. <https://doi.org/10.1016/j.nlm.2015.11.005>

Schnyer, D. M., Dobbins, I. G., Nicholls, L., Schacter, D. L., & Verfaellie, M. (2006). Rapid response learning in amnesia: Delineating associative learning components in repetition priming. *Neuropsychologia*, 44(1), 140–149. <https://doi.org/10.1016/j.neuropsychologia.2005.03.027>

Schomaker, J. (2019). Unexplored territory: Beneficial effects of novelty on memory. *Neurobiology of Learning and Memory*, 161, 46–50. <https://doi.org/10.1016/j.nlm.2019.03.005>

Schomaker, J., & Meeter, M. (2015). Short- and long-lasting consequences of novelty, deviance and surprise on brain and cognition. *Neuroscience & Biobehavioral Reviews*, 55, 268–279.  
<https://doi.org/10.1016/j.neubiorev.2015.05.002>

Schomaker, J., van Bronkhorst, M. L. V., & Meeter, M. (2014). Exploring a novel environment improves motivation and promotes recall of words. *Frontiers in Psychology*, 5.  
<https://doi.org/10.3389/fpsyg.2014.00918>

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, 91(6), 1402–1412.  
<https://doi.org/10.1016/j.neuron.2016.08.019>

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>

Shahbaba, B., Li, L., Agostinelli, F., Saraf, M., Cooper, K. W., Haghverdian, D., Elias, G. A., Baldi, P., & Fortin, N. J. (2022). Hippocampal ensembles represent sequential relationships among an extended sequence of nonspatial events. *Nature Communications*, 13(1), 787.

<https://doi.org/10.1038/s41467-022-28057-6>

Shamay-Tsoory, S. G., & Mendelsohn, A. (2019). Real-Life Neuroscience: An Ecological Approach to Brain and Behavior Research. *Perspectives on Psychological Science*, 14(5), 841–859.

<https://doi.org/10.1177/1745691619856350>

Sherman, B. E., & Turk-Browne, N. B. (2020). Statistical prediction of the future impairs episodic encoding of the present. *Proceedings of the National Academy of Sciences*, 117(37).

<https://doi.org/10.1073/pnas.2013291117/-DCSupplemental.y>

Sherman, B. E., Turk-Browne, N. B., & Goldfarb, E. V. (2024). Multiple Memory Subsystems: Reconsidering Memory in the Mind and Brain. *Perspectives on Psychological Science*, 19(1), 103–125. <https://doi.org/10.1177/17456916231179146>

Shin, Y. S., & DuBrow, S. (2021). Structuring Memory Through Inference-Based Event Segmentation. *Topics in Cognitive Science*, 13(1), 106–127. <https://doi.org/10.1111/tops.12505>

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

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., & Gluck, M. A. (2005). The role of dopamine in cognitive sequence learning: Evidence from Parkinson's disease. *Behavioural Brain Research*, 156(2), 191–199. <https://doi.org/10.1016/j.bbr.2004.05.023>

Shohamy, D., & Turk-Browne, N. B. (2013). Mechanisms for widespread hippocampal involvement in cognition. *Journal for Experimental Psychology: General*, 142(4), 1159–1170.

<https://doi.org/10.1037/a0034461>

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron*, 60(2), 378–389.

<https://doi.org/10.1016/j.neuron.2008.09.023>

Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes.

*Psychological Bulletin, 124*(3), 423–442. <https://doi.org/10.1037/0033-2909.124.3.423>

Silva, M., Baldassano, C., & Fuentemilla, L. (2019). Rapid memory reactivation at movie event boundaries promotes episodic encoding. *Journal of Neuroscience, 39*(43), 8538–8548.

<https://doi.org/10.1523/JNEUROSCI.0360-19.2019>

Simmons-Stern, N. R., Budson, A. E., & Ally, B. A. (2010). Music as a memory enhancer in patients with Alzheimer's disease. *Neuropsychologia, 48*(10), 3164–3167.

<https://doi.org/10.1016/j.neuropsychologia.2010.04.033>

Siquier, A., & Andrés, P. (2021). Episodic Memory Impairment in Parkinson's Disease: Disentangling the Role of Encoding and Retrieval. *Journal of the International Neuropsychological Society, 27*(3), 261–269. <https://doi.org/10.1017/S1355617720000909>

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical Methods*. Iowa State University Press.

Sommer, T. (2017). The Emergence of Knowledge and How it Supports the Memory for Novel Related Information. *Cerebral Cortex, 27*(3), 1906–1921. <https://doi.org/10.1093/cercor/bhw031>

Sommer, T., Hennies, N., Lewis, P. A., & Alink, A. (2022). The Assimilation of Novel Information into Schemata and Its Efficient Consolidation. *The Journal of Neuroscience, 42*(30), 5916–5929.

<https://doi.org/10.1523/JNEUROSCI.2373-21.2022>

Spalding, K. N., Jones, S. H., Duff, M. C., Tranel, D., & Warren, D. E. (2015). Investigating the Neural Correlates of Schemas: Ventromedial Prefrontal Cortex Is Necessary for Normal Schematic Influence on Memory. *Journal of Neuroscience, 35*(47), 15746–15751.

<https://doi.org/10.1523/JNEUROSCI.2767-15.2015>

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory*. Consulting Psychologists Press.

Spratling, M. W. (2016). Predictive coding as a model of cognition. *Cognitive Processing, 17*(3), 279–305. <https://doi.org/10.1007/s10339-016-0765-6>

- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93(24), 13515–13522.  
<https://doi.org/10.1073/pnas.93.24.13515>
- St. Jacques, P. L., & Schacter, D. L. (2013). Modifying Memory: Selectively Enhancing and Updating Personal Memories for a Museum Tour by Reactivating Them. *Psychological Science*, 24(4), 537–543. <https://doi.org/10.1177/0956797612457377>
- Stanley, D. (2022). apaTables: Create American Psychological Association (APA) Style Tables. *R Package Version 2.0.8.* <https://doi.org/10.5281/zenodo.683128>
- Stasiak, J. E., Mitchell, W. J., Reisman, S. S., Gregory, D. F., Murty, V. P., & Helion, C. (2023). Physiological arousal guides situational appraisals and metacognitive recall for naturalistic experiences. *Neuropsychologia*, 180, 108467.  
<https://doi.org/10.1016/j.neuropsychologia.2023.108467>
- Stawarczyk, D., Bezdek, M. A., & Zacks, J. M. (2021). Event Representations and Predictive Processing: The Role of the Midline Default Network Core. *Topics in Cognitive Science*, 13(1), 164–186.  
<https://doi.org/10.1111/tops.12450>
- Takeuchi, T., Tamura, M., Tse, D., Kajii, Y., Fernández, G., & Morris, R. G. M. (2022). Brain region networks for the assimilation of new associative memory into a schema. *Molecular Brain*, 15(1), 24. <https://doi.org/10.1186/s13041-022-00908-9>
- Talarico, J. M., & Rubin, D. C. (2003). Confidence, Not Consistency, Characterizes Flashbulb Memories. *Psychological Science*, 14(5), 455–461. <https://doi.org/10.1111/1467-9280.02453>
- Tambini, A., & Davachi, L. (2013). Persistence of hippocampal multivoxel patterns into postencoding rest is related to memory. *Proceedings of the National Academy of Sciences*, 110(48), 19591–19596. <https://doi.org/10.1073/pnas.1308499110>
- 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>

- Thibault, S., Py, R., Gervasi, A. M., Salemme, R., Koun, E., Lövden, M., Boulenger, V., Roy, A. C., & Brozzoli, C. (2021). Tool use and language share syntactic processes and neural patterns in the basal ganglia. *Science*, 374(6569), eabe0874. <https://doi.org/10.1126/science.abe0874>
- Tompany, A., & Thompson-Schill, S. L. (2021). Semantic influences on episodic memory distortions. *Journal of Experimental Psychology: General*, 150(9), 1800. <https://doi.org/10.1037/xge0001017>
- Torchiano, M. (2020). Effsize: A package for efficient effect size computation. *R Package Version 0.8.1*. <https://doi.org/10.5281/zenodo.683128>
- 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, 57–62. <https://doi.org/10.1038/s41586-018-0459-6>
- Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., Witter, M. P., & Morris, R. G. M. (2007). Schemas and memory consolidation. *Science*, 316(5821), 76–82. <https://doi.org/10.1126/science.1135935>
- 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>
- Tubridy, S., & Davachi, L. (2011). Medial Temporal Lobe Contributions to Episodic Sequence Encoding. *Cerebral Cortex*, 21(2), 272–280. <https://doi.org/10.1093/cercor/bhq092>
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology / Psychologie Canadienne*, 26(1), 1–12. <https://doi.org/10.1037/h0080017>
- Tulving, E., & Kroll, N. (1995). Novelty assessment in the brain and long-term memory encoding. *Psychonomic Bulletin & Review*, 2(3), 387–390. <https://doi.org/10.3758/BF03210977>
- Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *Hippocampus*, 8(3), 198–204. [https://doi.org/10.1002/\(SICI\)1098-1063\(1998\)8:3<198::AID-HIPO2>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1098-1063(1998)8:3<198::AID-HIPO2>3.0.CO;2-G)
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010).

- N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging*, 29(6), 1310–1320. IEEE Transactions on Medical Imaging. <https://doi.org/10.1109/TMI.2010.2046908>
- Umanath, S., & Marsh, E. J. (2014). Understanding How Prior Knowledge Influences Memory in Older Adults. *Perspectives on Psychological Science*, 9(4), 408–426.  
<https://doi.org/10.1177/1745691614535933>
- van Buuren, M., Kroes, M. C. W., Wagner, I. C., Genzel, L., Morris, X. G. M., & Fernández, G. (2014). Initial Investigation of the Effects of an Experimentally Learned Schema on Spatial Associative Memory in Humans. *Journal of Neuroscience*. <https://doi.org/10.1523/JNEUROSCI.2365-14.2014>
- van Halteren-van Tilborg, I. A. D. A., Scherder, E. J. A., & Hulstijn, W. (2007). Motor-Skill Learning in Alzheimer's Disease: A Review with an Eye to the Clinical Practice. *Neuropsychology Review*, 17(3), 203–212. <https://doi.org/10.1007/s11065-007-9030-1>
- van Kesteren, M. T. R., Beul, S. F., Takashima, A., Henson, R. N., Ruiter, D. J., & Fernández, G. (2013). Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent. *Neuropsychologia*, 51(12), 2352–2359.  
<https://doi.org/10.1016/j.neuropsychologia.2013.05.027>
- van Kesteren, M. T. R., Fernandez, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences*, 107(16), 7550–7555.  
<https://doi.org/10.1073/pnas.0914892107>
- van Kesteren, M. T. R., Krabbendam, L., & Meeter, M. (2018). Integrating educational knowledge: Reactivation of prior knowledge during educational learning enhances memory integration. *Npj Science of Learning*, 3(1), 1–8. <https://doi.org/10.1038/s41539-018-0027-8>
- van Kesteren, M. T. R., Rignanese, P., Gianferrara, P. G., Krabbendam, L., & Meeter, M. (2020). Congruency and reactivation aid memory integration through reinstatement of prior knowledge. *Scientific Reports*, 10(1), 1–13. <https://doi.org/10.1038/s41598-020-61737-1>

- van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., & Fernández, G. (2013). Consolidation Differentially Modulates Schema Effects on Memory for Items and Associations. *PLoS One*, 8(2).  
<https://doi.org/10.1371/journal.pone.0056155>
- van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., Morris, R. G. M., & Fernández, G. (2014). Building on prior knowledge: Schema-dependent encoding processes relate to academic performance. *Journal of Cognitive Neuroscience*, 26(10), 2250–2261. [https://doi.org/10.1162/jocn\\_a\\_00630](https://doi.org/10.1162/jocn_a_00630)
- 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>
- van Kesteren, M. T. R. van, Rijpkema, M., Ruiter, D. J., & Fernández, G. (2010). Retrieval of Associative Information Congruent with Prior Knowledge Is Related to Increased Medial Prefrontal Activity and Connectivity. *Journal of Neuroscience*, 30(47), 15888–15894.  
<https://doi.org/10.1523/JNEUROSCI.2674-10.2010>
- Virk, T., Letendre, T., & Pathman, T. (2024). The convergence of naturalistic paradigms and cognitive neuroscience methods to investigate memory and its development. *Neuropsychologia*, 196, 108779. <https://doi.org/10.1016/j.neuropsychologia.2023.108779>
- Wagner, I. C., van Buuren, M., Kroes, M. C., Gutteling, T. P., van der Linden, M., Morris, R. G., & Fernández, G. (2015). Schematic memory components converge within angular gyrus during retrieval. *eLife*, 4, 1–28. <https://doi.org/10.7554/elife.09668>
- Wang, S.-H., Redondo, R. L., & Morris, R. G. M. (2010). Relevance of synaptic tagging and capture to the persistence of long-term potentiation and everyday spatial memory. *Proceedings of the National Academy of Sciences*, 107(45), 19537–19542. <https://doi.org/10.1073/pnas.1008638107>
- Warren, D. E., Jones, S. H., Duff, M. C., & Tranel, D. (2014). False recall is reduced by damage to the ventromedial prefrontal cortex: Implications for understanding the neural correlates of schematic memory. *Journal of Neuroscience*, 34(22), 7677–7682. Scopus.  
<https://doi.org/10.1523/JNEUROSCI.0119-14.2014>

- Watson, D., Anna, L., & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales. *Journal of Personality and Social Psychology*.
- Webb, C. E., Turney, I. C., & Dennis, N. A. (2016). What's the gist? The influence of schemas on the neural correlates underlying true and false memories. *Neuropsychologia*, 93, 61–75.
- <https://doi.org/10.1016/j.neuropsychologia.2016.09.023>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wiestler, T., & Diedrichsen, J. (2013). Skill learning strengthens cortical representations of motor sequences. *eLife*, 2, e00801. <https://doi.org/10.7554/eLife.00801>
- Winocur, G., & Moscovitch, M. (2011). Memory transformation and systems consolidation. *Journal of the International Neuropsychological Society*, 17(5), 766–780.
- <https://doi.org/10.1017/S1355617711000683>
- Winograd, E., & Neisser, U. (2006). *Affect and Accuracy in Recall: Studies of “Flashbulb” Memories*. Cambridge University Press.
- Wittkuhn, L., & Schuck, N. W. (2021). Dynamics of fMRI patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nature Communications*, 12(1), Article 1.
- <https://doi.org/10.1038/s41467-021-21970-2>
- Woollett, K., & Maguire, E. A. (2010). The effect of navigational expertise on wayfinding in new environments. *Journal of Environmental Psychology*, 30(4), 565–573.
- <https://doi.org/10.1016/j.jenvp.2010.03.003>
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *NeuroImage*, 14(6), 1370–1386.
- <https://doi.org/10.1006/nimg.2001.0931>
- Wright, D. B. (1993). Recall of the Hillsborough disaster over time: Systematic biases of ‘flashbulb’

memories. *Applied Cognitive Psychology*, 7(2), 129–138.

<https://doi.org/10.1002/acp.2350070205>

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>

Yebra, M., Galarza-Vallejo, A., Soto-Leon, V., Gonzalez-Rosa, J. J., de Berker, A. O., Bestmann, S.,

Oliviero, A., Kroes, M. C. W., & Strange, B. A. (2019). Action boosts episodic memory encoding in humans via engagement of a noradrenergic system. *Nature Communications*, 10(1), Article 1.

<https://doi.org/10.1038/s41467-019-11358-8>

Yokoi, A., & Diedrichsen, J. (2019). Neural Organization of Hierarchical Motor Sequence

Representations in the Human Neocortex. *Neuron*, 103(6), 1178-1190.e7.

<https://doi.org/10.1016/j.neuron.2019.06.017>

Yousif, S. R., Lee, S. H., Sherman, B. E., & Papafragou, A. (2024). Event representation at the scale of ordinary experience. *Cognition*, 249, 105833. <https://doi.org/10.1016/j.cognition.2024.105833>

Zacks, J. M. (2020). Event Perception and Memory. *Annual Review of Psychology*, 71(Volume 71, 2020), 165–191. <https://doi.org/10.1146/annurev-psych-010419-051101>

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, 133(2), 273–293. <https://doi.org/10.1037/0033-2909.133.2.273>

Zadbood, A., Chen, J., Leong, Y. C., Norman, K. A., & Hasson, U. (2017). How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cerebral Cortex*, 27(10), 4988–5000. <https://doi.org/10.1093/cercor/bhw202>

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1), 45–57. IEEE Transactions on Medical Imaging.  
<https://doi.org/10.1109/42.906424>

Zheng, L., Gao, Z., McAvan, A. S., Isham, E. A., & Ekstrom, A. D. (2021). Partially overlapping spatial

environments trigger reinstatement in hippocampus and schema representations in prefrontal cortex. *Nature Communications*, 12(1), 6231. <https://doi.org/10.1038/s41467-021-26560-w>

Zion, D. B., Nevat, M., Prior, A., & Bitan, T. (2019). Prior Knowledge Predicts Early Consolidation in Second Language Learning. *Frontiers in Psychology*, 10.

<https://doi.org/10.3389/fpsyg.2019.02312>

## Appendix A: Chapter 1 Supplement

### A.1 Supplemental Methods

#### A.1.1 Bayesian multilevel regression models

When exploring the effect of event condition (predictable versus random) on temporal order memory, we also accounted for the influence of other relevant variables likely to interact with memory performance: stimulus category (animals vs. foods), block number (1-8), event position within a list (1-4), and sequence position within an event (1-6). Specifically, we ran a series of Bayesian multilevel logistic regression models (one per experiment), in which binary accuracy on each memory test trial was predicted by condition plus all confound variables. Categorical predictors were effect-coded and continuous predictors were mean-centered prior to analysis. All models included a random intercept for each subject, as well as both fixed and random slopes for all predictor variables. Models were run using the *brms* package (ver. 2.10.0; Bürkner, 2017) in R. All models included uninformative priors for all parameters and were fit using 4 Markov chains, each with 8000 iterations and 1000 warmup iterations.

We used Bayesian regression for this control analysis because it more readily allowed us to estimate complex multilevel models with numerous random effects (whereas comparable Frequentist models produced convergence errors). Critically, when using uninformative priors (as we do here), Frequentist and Bayesian approaches can be treated as conceptually analogous. Although Bayesian models do not yield traditional p-values, we can assess the “significance” of all fixed effects by examining whether the 95% credibility interval ( $95\%-CrI$ ) of the corresponding beta coefficient contains zero. A lack of such overlap is taken as evidence that the effect is reliably different from zero.

## A.2 Supplemental Results

### A.2.1 Controlling for confounds in the effect of condition on order memory

For each of our three experiments, we ran a Bayesian logistic regression model (see Supplemental Methods) to examine the effect of condition (predictable vs. random) on trial-level accuracy during the order reconstruction test after controlling for a set of confound variables: stimulus category (animals vs. foods), block number (1-8), event position within a list (1-4), and sequence position within an event (1-6). In each of these three Bayesian models (one for each experiment), controlling for these variables did not change our finding of enhanced order memory for items seen in the predictable versus random store. That is, the effect of condition on memory remained robust (i.e., credibility intervals did not contain zero) in Experiment 1 ( $b_{cond} = 0.54$ ,  $SE = 0.11$ ,  $95\%-CrI = [0.32, 0.76]$ ), Experiment 2 ( $b_{cond} = 0.28$ ,  $SE = 0.12$ ,  $95\%-CrI = [0.06, 0.51]$ ), and Experiment 3 ( $b_{cond} = 0.30$ ,  $SE = 0.12$ ,  $95\%-CrI = [0.06, 0.53]$ ). Full model outputs can be found in Supplemental Tables 1-3.

### A.2.2 Order memory and aisle response accuracy during encoding

In Experiments 1 and 3, participants were required to use their memory for the predictable store's aisle sequence to guide their behavior during encoding. In Chapter 1, we showed that participants were generally very accurate in making these responses, reaching a level of performance that was comparable to the random condition (in which cues were provided) after the first block. We also excluded all participants who did not make predictable aisle response with accuracy  $> 80\%$  from any memory test analyses, ensuring that all participants in our order memory analyses had demonstrated sufficient learning of the predictable action sequence (see Methods). Nevertheless, in relatively rare cases where participants failed to execute the correct aisle response, we cannot be sure that they had access to a representation of the learned action sequence while encountering novel items. To address this concern, we re-examined the difference in order memory (i.e., ordinal accuracy) across conditions including only items in which participants had executed the correct aisle response (prior to viewing the item) during encoding. We found that order memory remained significantly greater for items from the predictable versus random

store in both Experiment 1 ( $t(62) = 4.55, p < 0.001, CI = [0.05, 0.14], d = 0.48, CI_d = [0.26, 0.70]$ ) and Experiment 3 ( $t(65) = 2.72, p = 0.008, CI = [0.01, 0.09], d = 0.26, CI_d = [0.07, 0.46]$ ).

### A.2.3 Order memory and aisle repetition during encoding

As discussed in the Methods, all 6-item aisle sequences were structured such that 2 out of 4 aisles were visited twice per sequence, and the remaining 2 aisles were visited only once. Because of this structure, it is possible that aisles visited twice per sequence (“repeated aisles”) were more likely to produce interference in memory (i.e., between distinct items that share the same location). Both predictable and random sequences had the same number of repeated aisles, making it unlikely that any potential interference produced by such aisles would explain our order memory effects. Nevertheless, to directly rule out this possibility, we examined ordinal accuracy separately for items paired with repeated versus non-repeated aisles. In all three experiments, we found that ordinal accuracy was greater for items from predictable versus random events, both when the aisle associated with each item was visited twice during the event (Exp. 1:  $t(62) = 4.31, p < 0.001, CI = [0.05, 0.15], d = 0.49, CI_d = [0.25, 0.72]$ ; Exp 2:  $t(60) = 2.62, p = 0.011, CI = [0.01, 0.11], d = 0.28, CI_d = [0.06, 0.50]$ ; Exp 3:  $t(65) = 2.38, p = 0.020, CI = [0.01, 0.09], d = 0.24, CI_d = [0.04, 0.44]$ ), and when it was not (Exp. 1:  $t(62) = 3.52, p < 0.001, CI = [0.03, 0.12], d = 0.37, CI_d = [0.16, 0.59]$ ; Exp 2:  $t(60) = 2.33, p = 0.023, CI = [0.01, 0.10], d = 0.27, CI_d = [0.04, 0.50]$ ; Exp 3:  $t(65) = 2.50, p = 0.015, CI = [0.01, 0.09], d = 0.25, CI_d = [0.05, 0.44]$ ).

We also ran a series of logistic regression models in which ordinal accuracy on each trial was predicted by condition (predictable vs. random), whether each aisle was repeated or not, and sequence position. Sequence position was included as an additional covariate in this model because it is partly confounded with aisle repetition, in that the first aisle within an event is never revisited (see Methods). These models revealed that repeated aisles were associated with worse ordinal accuracy (Exp. 1:  $b_{rep} = -0.18, SE = 0.07, CI = [-0.31, -0.05], p = 0.006$ ; Exp 2:  $b_{rep} = -0.31, SE = 0.07, CI = [-0.44, -0.18], p < 0.001$ ; Exp 3:  $b_{rep} = -0.19, SE = 0.07, CI = [-0.33, -0.06], p = 0.004$ ), consistent with the idea that associating two items with the same aisle/spatial location might generate interference in temporal memory. Critically, however, the effect of condition remained strongly significant even when controlling

for aisle repetition (Exp. 1:  $b_{cond} = 0.49$ ,  $SE = 0.11$ ,  $CI = [0.28, 0.71]$ ,  $p < 0.001$ ; Exp 2:  $b_{cond} = 0.29$ ,  $SE = 0.11$ ,  $CI = [0.08, 0.50]$ ,  $p = 0.006$ ; Exp 3:  $b_{cond} = 0.29$ ,  $SE = 0.11$ ,  $CI = [0.07, 0.50]$ ,  $p = 0.009$ ). Earlier sequence positions were also associated with better ordinal accuracy, as also revealed by the logistic regression models discussed in Chapter 1 (Exp. 1:  $b_{seq\_pos} = -0.23$ ,  $SE = 0.02$ ,  $CI = [-0.27, -0.18]$ ,  $p < 0.001$ ; Exp 2:  $b_{seq\_pos} = -0.22$ ,  $SE = 0.03$ ,  $CI = [-0.27, -0.17]$ ,  $p < 0.001$ ; Exp 3:  $b_{seq\_pos} = -0.19$ ,  $SE = 0.02$ ,  $CI = [-0.23, -0.14]$ ,  $p < 0.001$ ).

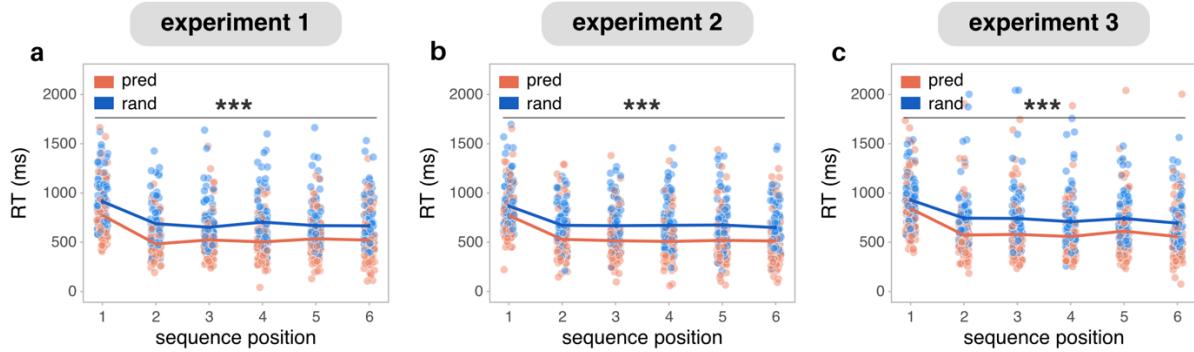
#### A.2.4 Spatial memory performance and aisle location

In Experiments 1 and 2, we examined how spatial memory performance differed as a function of aisle location. In Experiment 1, a repeated measures ANOVA revealed that participants' spatial memory performance varied across locations ( $F(3, 186) = 6.67$ ,  $p < 0.001$ , partial  $\eta^2 = 0.10$ ,  $CI_\eta = [0.02, 0.17]$ ), with memory for items at the leftmost edge of the screen (aisle 1) being better than all other aisles (aisles 2-4) (aisle 1 vs. 2:  $t(62) = 2.89$ ,  $p_{FDR} = 0.011$ ,  $CI_{FCR} = [0.01, 0.10]$ ,  $d = 0.30$ ,  $CI_d = [0.09, 0.50]$ ; aisle 1 vs. 3:  $t(62) = 3.69$ ,  $p_{FDR} = 0.001$ ,  $CI_{FCR} = [0.02, 0.11]$ ,  $d = 0.35$ ,  $CI_d = [0.16, 0.55]$ ; aisle 1 vs. 4:  $t(62)$ ,  $p_{FDR} = 0.001$ ,  $CI_{FCR} = [0.02, 0.10]$ ,  $d = 0.32$ ,  $CI_d = [0.15, 0.48]$ ). A similar effect was present in Experiment 2 ( $F(2.45, 146.97) = 13.868$ ,  $p < 0.001$ , partial  $\eta^2 = 0.19$ ,  $CI_\eta = [0.08, 0.29]$ ; Greenhouse-Geisser correction for non-sphericity applied), such that memory was again best for items that appeared in the leftmost aisle (aisle 1 vs. 2:  $t(60) = 5.13$ ,  $p_{FDR} < 0.001$ ,  $CI_{FCR} = [0.05, 0.11]$ ,  $d = 0.45$ ,  $CI_d = [0.27, 0.63]$ ; aisle 1 vs. 3:  $t(60) = 5.73$ ,  $p_{FDR} < 0.001$ ,  $CI_{FCR} = [0.06, 0.12]$ ,  $d = 0.51$ ,  $CI_d = [0.32, 0.70]$ ; aisle 1 vs. 4:  $t(60) = 3.49$ ,  $p_{FDR} = 0.002$ ,  $CI_{FCR} = [0.02, 0.07]$ ,  $d = 0.21$ ,  $CI_d = [0.09, 0.33]$ ). In this sample, spatial memory for the items in the rightmost aisle was also better than those in the middle aisles (aisle 4 vs. 2:  $t(60) = 2.11$ ,  $p_{FDR} = 0.047$ ,  $CI_{FCR} = [0.0005, 0.07]$ ,  $d = 0.18$ ,  $CI_d = [0.01, 0.35]$ ; aisle 4 vs. 3:  $t(60) = 2.61$ ,  $p_{FDR} = 0.017$ ,  $CI_{FCR} = [0.01, 0.08]$ ,  $d = 0.23$ ,  $CI_d = [0.05, 0.41]$ ). These effects complicate the interpretation of our finding that participants' spatial memory for the first item in an event was enhanced in the predictable versus random store, given that the location of these first items was not balanced across conditions.

### A.2.5 Reminder task performance

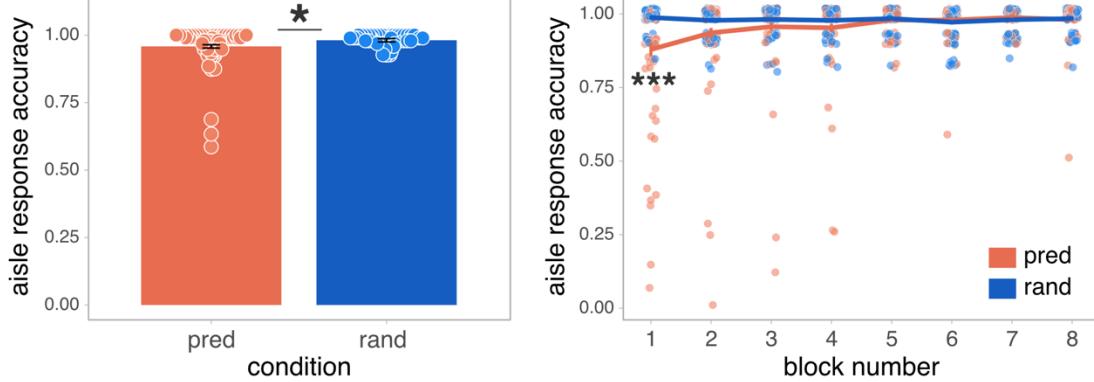
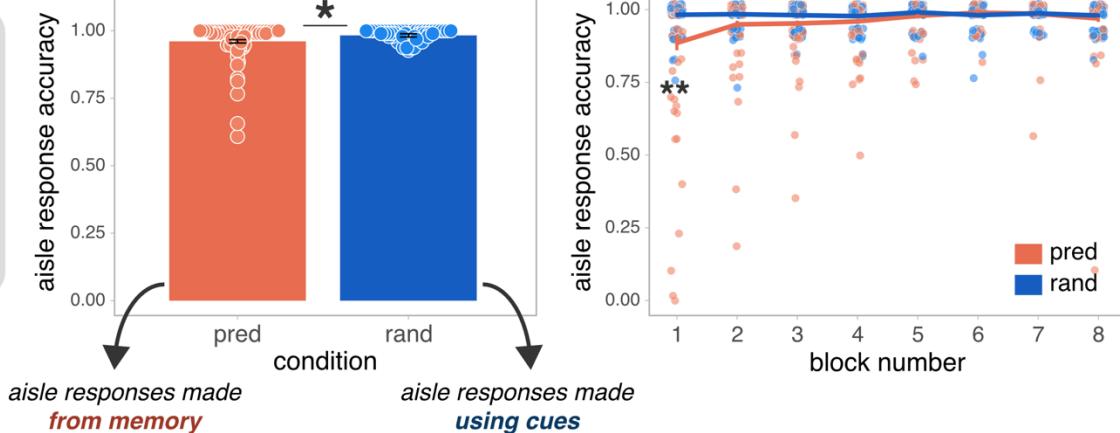
In Experiments 1 and 3, participants were required to make aisle responses in the predictable store from memory while seeing cues in the random store. Because of this memory demand, we included an explicit reminder of the predictable sequence halfway through the experiment (i.e., after the fourth retrieval block; Fig. 1a). Although the high accuracy of aisle responses throughout encoding (Supplemental Fig. 2) clearly indicates that participants retained memory for the action sequence, we can further confirm this conclusion by checking behavior on the reminder task. Participants began this task by passively observing the learned aisle sequence twice, before then recreating it from memory twice during two consecutive test trials. As expected, performance during these test trials was near ceiling in both Experiment 1 ( $M = 0.99$ ,  $SD = 0.05$ ) and Experiment 3 ( $M = 0.99$ ,  $SD = 0.04$ ).

### A.3 Supplemental Figures

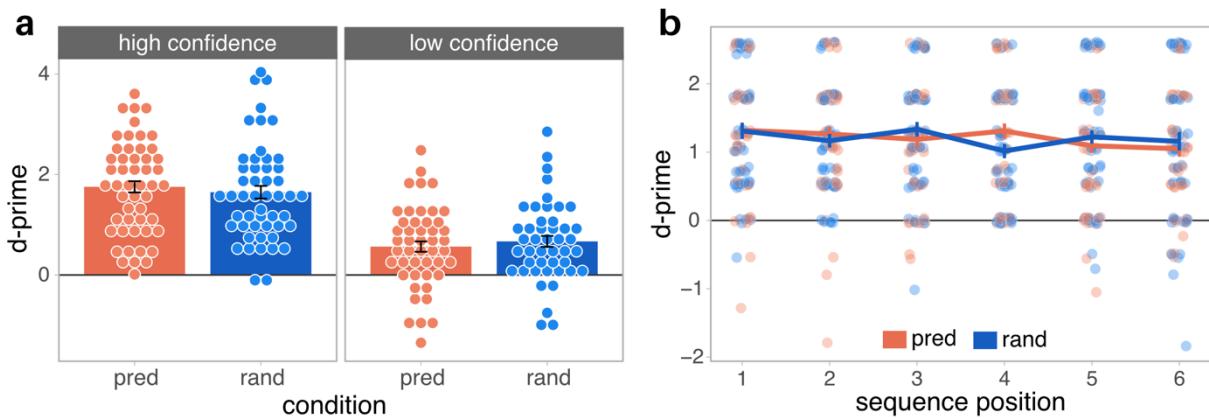


**Appendix A, Figure 1. Reaction times (RTs) to aisle cues during the intermixed pretraining study block.**

Participants' speed in executing cued aisle responses across conditions (predictable versus random) and sequence positions, for Experiment 1 (a), Experiment 2 (b), and Experiment 3 (c). Statistical significance was assessed using Wilcoxon signed-rank tests, FDR-corrected for multiple comparisons. Note that although these plots depict data from all participants included in pretraining analyses, statistics were based on N = 76, N = 77, and N = 72 participants for Experiments 1-3 respectively, given that a small set of participants did not have responses in each condition and sequence position bin after trial-level exclusions (see Methods). Error bars indicate within-subject standard errors, and dots represent individual participants. \*\*\* p < 0.001.

**a****experiment 1****b****experiment 3**

**Appendix A, Figure 2. Accuracy of participants' aisle responses during encoding.** Encoding data from Experiment 1 (a) and 3 (b). Left plots show the difference in accuracy averaged across all blocks. Right plots show accuracy in both conditions as a function of block. Statistical significance was assessed using paired, two-tailed t-tests, with FDR-correction for multiple comparisons performed for right plots. Error bars indicate within-subject standard errors, and dots represent individual participants. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .



**Appendix A, Figure 3. Item recognition by confidence and sequence position for Experiment 3.** D-prime scores (i.e., item recognition accuracy) broken down by high versus low confidence bins **(a)** and by sequence position **(b)**. There were no statistically significant differences between predictable and random d-prime scores within any confidence or sequence position bins, according to paired, two-tailed t-tests, with FDR correction applied to comparisons in **(b)**. Error bars indicate within-subject standard errors, and dots represent individual participants.

## A.4 Supplemental Tables

**Appendix A, Table 1.** Experiment 1: Results of a Bayesian, multilevel logistic regression model examining the effect of condition on order reconstruction memory after controlling for confounds. *CrI* = credibility interval.

	<i>Estimate</i>	<i>Est. Error</i>	<i>l-95% CrI</i>	<i>u-95% CrI</i>
<b>Fixed Effects</b>				
Intercept	0.480	0.123	0.242	0.722
condition	0.539	0.113	0.317	0.761
seq position	-0.252	0.024	-0.299	-0.205
event position	-0.086	0.041	-0.167	-0.007
block	-0.006	0.026	-0.058	0.045
stim cat	0.354	0.111	0.139	0.572
<b>Random Effects</b>				
sd(Intercept)	0.941	0.096	0.772	1.147
sd(condition)	0.441	0.239	0.021	0.837
sd(seq position)	0.119	0.029	0.061	0.176
sd(event position)	0.222	0.044	0.139	0.312
sd(block)	0.170	0.023	0.128	0.219
sd(stim cat)	0.485	0.233	0.034	0.860

**Appendix A, Table 2.** Experiment 2: Results of a Bayesian, multilevel logistic regression model examining the effect of condition on order reconstruction memory after controlling for confounds. *CrI* = credibility interval.

	<i>Estimate</i>	<i>Est. Error</i>	<i>l-95% CrI</i>	<i>u-95% CrI</i>
<b>Fixed Effects</b>				
Intercept	0.524	0.141	0.250	0.806
condition	0.283	0.115	0.059	0.509
seq position	-0.253	0.026	-0.305	-0.202
event position	-0.078	0.043	-0.163	0.008
block	-0.060	0.024	-0.107	-0.014
stim cat	0.203	0.115	-0.025	0.428
<b>Random Effects</b>				
sd(Intercept)	1.059	0.111	0.864	1.296
sd(condition)	0.469	0.240	0.027	0.865
sd(seq position)	0.141	0.027	0.091	0.197
sd(event position)	0.247	0.043	0.166	0.337
sd(block)	0.142	0.022	0.101	0.188
sd(stim cat)	0.480	0.240	0.029	0.872

**Appendix A, Table 3.** Experiment 3: Results of a Bayesian, multilevel logistic regression model examining the effect of condition on order reconstruction memory after controlling for confounds. *CrI* = credibility interval.

	<i>Estimate</i>	<i>Est. Error</i>	<i>l-95% CrI</i>	<i>u-95% CrI</i>
<b>Fixed Effects</b>				
Intercept	1.037	0.152	0.745	1.341
condition	0.298	0.120	0.060	0.531
seq position	-0.217	0.023	-0.263	-0.171
event position	-0.070	0.045	-0.158	0.016
block	0.040	0.029	-0.017	0.097
stim cat	0.160	0.121	-0.081	0.399
<b>Random Effects</b>				
sd(Intercept)	1.178	0.121	0.967	1.445
sd(condition)	0.513	0.264	0.031	0.940
sd(seq position)	0.107	0.030	0.044	0.164
sd(event position)	0.269	0.045	0.185	0.362
sd(block)	0.201	0.026	0.155	0.256
sd(stim cat)	0.517	0.265	0.029	0.950

**Appendix A, Table 4.** Comparison of order reconstruction memory and spatial memory effects across experiments. Statistics come from independent two-sample t-tests; assumption of equal variance verified using F-tests (Snedecor & Cochran, 1989). *CI* = confidence interval for *t* estimate. *d* = Cohen's *d*. *CI<sub>d</sub>* = confidence interval for *d* estimate.

	<i>t</i>	<i>df</i>	<i>p</i>	95% <i>CI</i>	<i>d</i>	95% <i>CI<sub>d</sub></i>
<b>Order memory</b>						
Exp. 1 vs. Exp. 2	1.175	122	0.242	[-0.024, 0.094]	0.211	[-0.15, 0.57]
Exp. 1 vs. Exp. 3	1.531	127	0.128	[-0.013, 0.098]	0.270	[-0.08, 0.62]
Exp. 2 vs. Exp. 3	0.298	125	0.780	[-0.048, 0.064]	0.050	[-0.30, 0.40]
<b>Spatial memory</b>						
Exp. 1 vs. Exp. 2	-1.348	122	0.180	[-0.083, 0.016]	0.242	[-0.11, 0.60]

## Appendix B: Chapter 2 Supplement

### B.1 Supplemental Methods & Results

#### Bayesian multilevel regression models

As described in the *Methods* of Chapter 2, we began all linear regression analyses by running Frequentist multilevel models with random intercepts and random slopes for each main effect (all nested within participants). However, when Frequentist models with this full set of random effects did not properly converge, we instead used Bayesian estimation methods, which better allowed for reliable estimation of the complex hierarchical structure within many of our regression models. Importantly, when using uninformative priors, as we do here, Bayesian models can be thought of as conceptually analogous to their Frequentist counterparts.

All Bayesian models were implemented using the *brms* package in R (ver. 2.10.0; Bürkner, 2017) in R. Uninformative (i.e., flat) priors were used for all parameters. Each model was fit using 4 Markov chains, with between 8000 and 10000 iterations per chain (such that the number of iterations was increased until all parameters for the given model converged). As with all Frequentist models used throughout our analyses, all categorical predictors were effect-coded and continuous predictors were mean-centered. Although Bayesian models do not produce p-values, the statistical reliability of fixed effects can be assessed by examining whether a parameter's 95% credibility interval (*CrI*) contains zero. A lack of such overlap is taken as evidence that the effect is reliably different from zero. We also report directional posterior probabilities (*pd*) for each fixed effect. These probabilities reflect the proportion of a parameter's posterior distribution that is in the same direction as the fixed effect estimate (i.e., positive or negative). Similar to the relationship between traditional p-values and confidence intervals in a Frequentist framework, when a parameter's 95% credibility interval excludes 0, the directional posterior probability will exceed 0.975.

#### Template activation searchlight

To explore the degree to which information about the predictable action sequence was

represented throughout the brain — beyond our *a priori* ROIs — we conducted an exploratory whole-brain searchlight analysis. First, because this type of analysis requires that data from all participants aligned to the same space, we reran the GLMs created to extract brain activity patterns from the sequence localizer and object sequence encoding tasks in MNI space (instead of subject-specific T1 space). We then searched for evidence of predictable template activation in cubic searchlights tiled throughout the brain, each with a side length of 7 voxels. Searchlights were centered on all voxels that fell within the union of the neocortical and hippocampal masks from FSL’s Harvard-Oxford Structural atlas. If any participant was missing data (due to BOLD signal dropout) from more than 50% of the voxels within a given searchlight cube, or if over 50% of voxels fell outside of the brain mask, that cube was excluded from the analysis.

Within each cubic searchlight, we followed the same general sequence of steps to compute template activation as in our ROI-based analysis: (1) defining predictable and variable sequence templates from the sequence localizer task; (2) computing the representational similarity between these templates and activity patterns evoked during all encoding events; (3) contrasting the strength of predictable template activation during predictable encoding events (P) with our set of three control measures (C1: predictable template activation during variable events; C2: variable template activation during variable events; and C3: variable template activation during predictable events; see Chapter 2 Methods for details). Specifically, within each cube, we computed the within-subject difference between predictable template activation (P) and each control measure (C1-3) using a set of three paired t-tests. In order to quantify the degree to which predictable template activation surpassed all three controls, we first calculated the mean t-statistic across these tests. That mean value was then converted to a z-statistic and assigned to the center voxel of the current cube. After repeating this procedure for each voxel/searchlight cube, we used FSL’s *cluster* tool to perform cluster-level statistical inference on the resulting whole-brain volume using Gaussian Random Field (GRF) theory, with a cluster-defining threshold of  $z = 2.3$  and a corrected significance level of  $p < .05$ . Critically, we then thresholded this cluster-corrected volume with a mask of voxels where all three *individual* control comparisons were significant (and in the expected direction).

This final masking step ensured that significant voxels in the searchlight volume did not reflect cases where one or more of the t-tests were non-significant, but the average statistic was not (e.g., if two of the three t-tests were strong enough to compensate for a non-significant third test).

Ultimately, the searchlight analysis revealed effects that were broadly consistent with our corresponding ROI-based analysis. We observed evidence for selective predictable template activation in clusters within both PMC and mPFC (at an uncorrected statistical threshold,  $p < .01$ ), as well as clusters overlapping with premotor cortex. Effects also emerged in primary motor cortex (M1), supplementary motor area (SMA), pre-SMA, parietal cortex, and lateral prefrontal cortex. Interestingly, several of these predictable template activation clusters — especially within the lateral prefrontal cortex — were largely right-lateralized. For a full depiction of searchlight results, see Appendix B, Figure 1.

### **Template activation across sequence positions**

Our main template activation analysis approach searches for representations of the predictable motor structure that are sustained across the full encoding event (i.e., as participants encode the full six-item object sequence). Although this approach reflects our hypothesis that this motor representation functions as a stable scaffold for incoming objects, it is nevertheless possible that its strength fluctuates throughout the event. For example, evidence for this scaffold may be most prevalent at the start of the event, when participants must first access their knowledge of the predictable action sequence.

To explore this possibility, we sought to quantify the strength of the predictable (and variable) sequence templates while each individual object in an event was being encoded. We used the activity patterns obtained for each individual sequence position as part of the “within-event stability” analysis described in Chapter 2. Briefly, these patterns were obtained using an LSS modeling approach (Mumford et al., 2012), whereby we constructed a separate GLM for each sequence position in each encoding event, resulting in 288 models (6 positions \* 4 events \* 12 blocks) per participant. Each model contained one regressor of interest corresponding to the presentation of a single object and its preceding motor/aisle response and feedback period, along with regressors of non-interest corresponding to all other sequence positions and task phases. From these GLMs, we obtained a set of statistical maps corresponding to brain

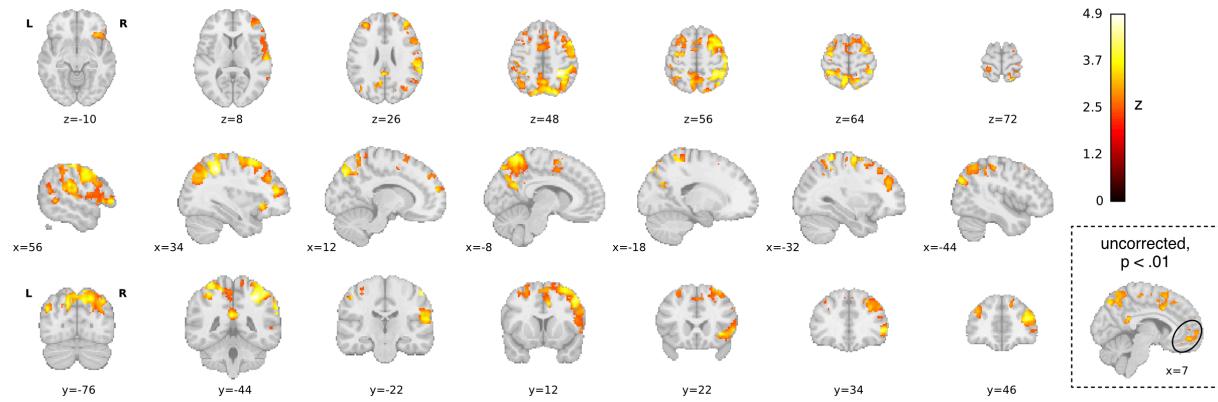
activity associated with each individual object during the encoding/errand task. From these maps, multivariate patterns were extracted within each ROI and then z-scored.

To assess predictable template activation at each sequence position, we then followed a procedure analogous to that described in Chapter 2. Specifically, we computed the representational similarity (i.e., Pearson correlation, with Fisher's R-to-Z transformation applied) between the predictable and random sequence templates (derived from our main template activation analysis) and each of these individual object encoding patterns. As described in the *Methods*, this analysis results in 4 measures of template activation per object/sequence position (Figure 2b). Our main measure of interest was similarity between the *predictable sequence template and encoding patterns from predictable events* (P). This measure of predictable template activation was then compared against three control measures: similarity between the *predictable sequence template and random encoding patterns* (C1); similarity between the *random sequence template and random encoding patterns* (C2); and similarity between the *random sequence template and predictable encoding patterns* (C3). To the extent that predictable template activation (P) exceeds these controls, we argue that a given brain region selectively represents information about the predictable motor structure during novel object encoding in the predictable store. For more details about the conceptual interpretation of each control measure, see the *Methods*.

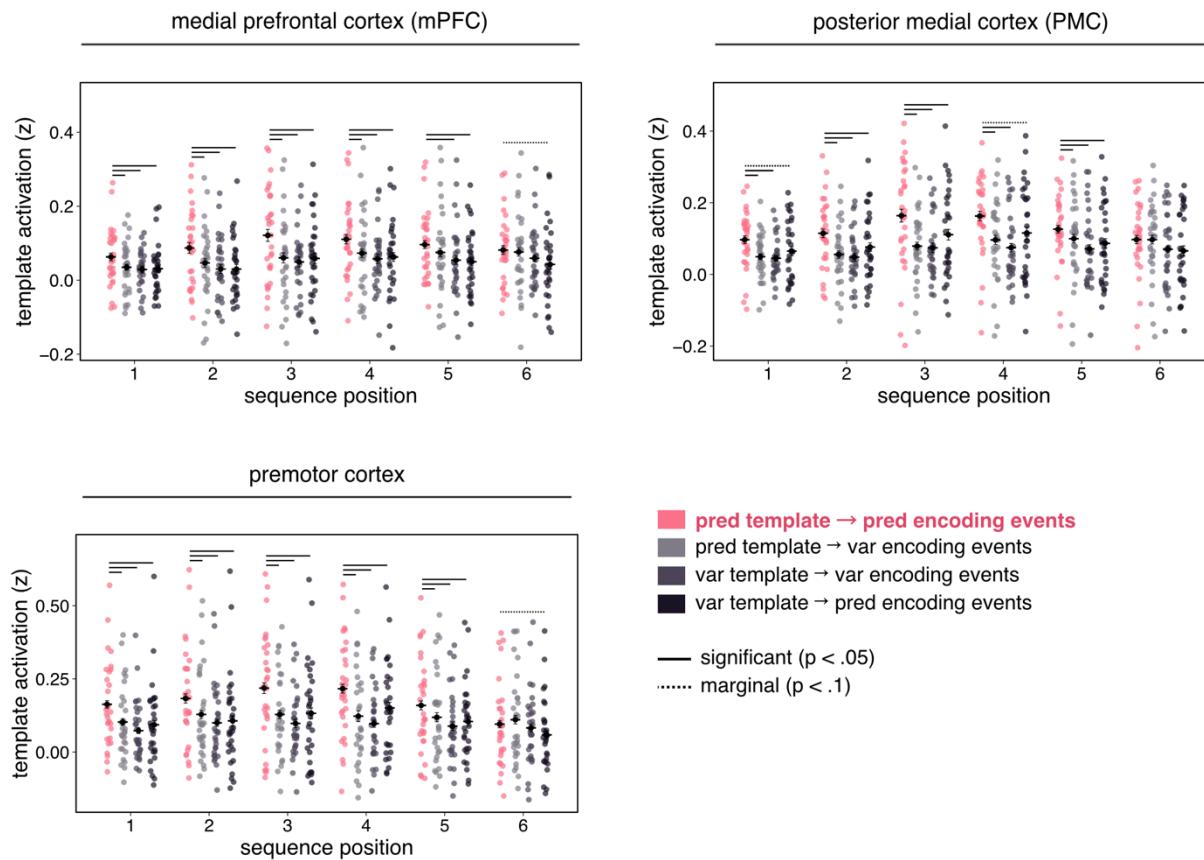
This analysis revealed that in all ROIs showing evidence for predictable template activation — namely, mPFC, PMC, and premotor cortex — the strength such activation was generally stable throughout the event, dropping off only at the final sequence position, when participants could presumably let the predictable action sequence drop from their working memory. See Appendix B, Figure 2 for a visualization of these effects.

## B.2 Supplemental Figures

**template activation searchlight**  
conjunction of all control comparisons



**Appendix B, Figure 1. Template activation whole-brain searchlight.** Depicts regions where predictable template activation (i.e., similarity between the predictable sequence template and predictable encoding patterns) exceeds all control measures. Results are corrected for multiple comparisons using Gaussian Random Field (GRF) theory with a cluster-defining threshold of  $z = 2.3$  and a corrected significance level of  $p < .05$  (except for the brain slice depicted in the box on the right, which shows results at an uncorrected significance level of  $p < .01$ . See Supplemental Methods & Results above for more details on how the searchlight volume was generated.



**Appendix B, Figure 2. Template activation across sequence positions.** Depicts similarity between sequence templates and patterns evoked during each individual sequence position within an encoding event. These encoding patterns reflect activity evoked during the period extending from each aisle response cue to the offset of the following object stimulus. Error bars indicate within-subjects standard errors. Solid lines indicate comparisons that are significant at  $p < .05$ , and dashed lines indicate marginal effects ( $p < .1$ ). Note that these significance values are not corrected for multiple comparisons.

### B.3 Supplemental Tables

**Appendix B, Table 1.** Results from Bayesian multilevel regression models in which order memory performance is predicted by condition, block, aisle response RTs, stimulus category, and the interaction between condition and block. Models include random slopes for all main effects, as well as random intercepts for each participant. Top and bottom panels represent models using different metrics of order accuracy: the proportion of objects selected in the correct ordinal position (*top*) and Levenshtein distance between the recalled and correct object sequences (*bottom*). Directional posterior probability (*pd*) refers to the proportion of each model term's posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd* > 0.975).

order memory ~ condition × block + stimulus category + aisle response RT (Bayesian)					
	Fixed Effect	Estimate	Std. Error	95% CrI	pd
<i>ordinal accuracy</i>					
	condition	0.038	0.019	[0.001, 0.077]	<b>0.978</b>
	block	0.016	0.003	[0.010, 0.022]	<b>1.000</b>
	stimulus category	-0.047	0.020	[-0.087, -0.007]	<b>0.989</b>
	aisle response RT (log-transformed)	-0.009	0.035	[-0.077, 0.060]	0.600
	condition × block	0.006	0.004	[-0.002, 0.013]	0.933
<i>Levenshtein distance</i>					
	condition	-0.179	0.087	[-0.357, -0.004]	<b>0.978</b>
	block	-0.094	0.015	[-0.124, -0.064]	<b>1.000</b>
	stimulus category	0.273	0.089	[0.089, 0.453]	<b>0.997</b>
	aisle response RT (log-transformed)	0.029	0.179	[-0.323, 0.383]	0.565
	condition × block	-0.035	0.019	[-0.073, 0.001]	0.971

*CrI* = credibility interval

*pd* = directional posterior probability

**Appendix B, Table 2.** Results from Bayesian multilevel regression models in which item recognition was predicted by condition, block, aisle response RTs, stimulus category, and the interaction between condition and block. Models include random slopes for all main effects, as well as random intercepts for each participant. Separate models were run for events tested on the same day as encoding (“immediate”, *top*) and 24 hours later (“delay”, *bottom*). Directional posterior probability (*pd*) refers to the proportion of each model term’s posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd* > 0.975).

item recognition ~ condition × block + stimulus category + aisle response RT (Bayesian)

	Fixed Effect	Estimate	Std. Error	95% CrI	pd
<i>immediate timepoint</i>					
	condition	-0.001	0.015	[-0.031, 0.028]	0.536
	block	0.002	0.002	[-0.002, 0.006]	0.880
	stimulus category	-0.099	0.015	[-0.128, -0.070]	<b>1.000</b>
	aisle response RT (log-transformed)	0.016	0.036	[-0.057, 0.088]	0.670
	condition × block	0.004	0.004	[-0.002, 0.011]	0.894
<i>delay timepoint</i>					
	condition	0.026	0.015	[-0.005, 0.056]	0.948
	block	0.003	0.002	[-0.001, 0.008]	0.923
	stimulus category	-0.054	0.016	[-0.086, -0.023]	> <b>0.999</b>
	aisle response RT (log-transformed)	0.033	0.037	[-0.038, 0.104]	0.817
	condition × block	-0.005	0.004	[-0.014, 0.003]	0.899

*CrI* = credibility interval

*pd* = directional posterior probability

**Appendix B, Table 3.** Results from Frequentist multilevel regression model examining the relationship between order memory performance and template activation across conditions and blocks. Models include random slopes for all main effects, as well as random intercepts for each participant.

order memory ~ template activation × condition × block					
	Fixed Effect	Estimate	Std. Error	95% CI	p
<b>mPFC</b>					
template activation	0.084	0.038	[0.002, 0.165]	<b>0.044</b>	
condition	0.031	0.020	[-0.009, 0.071]	0.126	
block	0.017	0.003	[0.012, 0.023]	< <b>0.001</b>	
template activation × condition	0.041	0.069	[-0.095, 0.177]	0.553	
template activation × block	0.033	0.009	[0.015, 0.052]	< <b>0.001</b>	
condition × block	0.003	0.004	[-0.004, 0.011]	0.386	
template activation × condition × block	0.019	0.018	[-0.015, 0.054]	0.276	
<b>PMC</b>					
template activation	0.039	0.041	[-0.046, 0.124]	0.355	
condition	0.037	0.020	[-0.004, 0.079]	0.075	
block	0.016	0.003	[0.011, 0.022]	< <b>0.001</b>	
template activation × condition	0.040	0.070	[-0.098, 0.178]	0.572	
template activation × block	0.042	0.009	[0.024, 0.060]	< <b>0.001</b>	
condition × block	0.002	0.004	[-0.006, 0.009]	0.670	
template activation × condition × block	0.040	0.018	[0.006, 0.075]	<b>0.022</b>	
<b>premotor cortex</b>					
template activation	-0.076	0.045	[-0.169, 0.017]	0.106	
condition	0.048	0.021	[0.005, 0.090]	<b>0.031</b>	
block	0.016	0.003	[0.010, 0.021]	< <b>0.001</b>	
template activation × condition	0.080	0.073	[-0.065, 0.225]	0.276	
template activation × block	0.028	0.009	[0.010, 0.047]	<b>0.003</b>	
condition × block	0.003	0.004	[-0.005, 0.011]	0.516	
template activation × condition × block	0.020	0.017	[-0.013, 0.054]	0.239	

**Appendix B, Table 4.** Results from Frequentist multilevel regression model examining the relationship between order memory and within-event stability across conditions and blocks. Models include random slopes for all main effects and random intercepts within each participant, unless denoted with †, in which case only random intercepts were included (due to model convergence issues). For these random intercept-only models, we report results from corresponding Bayesian models with random slopes in **Appendix B, Table 5**.

order memory ~ stability × condition × block					
	Fixed Effect	Estimate	Std. Error	95% CI	p
<b>mPFC</b>					
stability	0.302	0.066	[0.167, 0.437]	< <b>0.001</b>	
condition	0.031	0.021	[-0.012, 0.074]	0.153	
block	0.017	0.003	[0.011, 0.023]	< <b>0.001</b>	
stability × condition	-0.163	0.096	[-0.352, 0.026]	0.091	
stability × block	0.007	0.014	[-0.019, 0.034]	0.588	
condition × block	0.005	0.004	[-0.002, 0.013]	0.142	
stability × condition × block	0.048	0.026	[-0.003, 0.098]	0.063	
<b>PMC</b>					
stability	0.434	0.073	[0.284, 0.584]	< <b>0.001</b>	
condition	0.023	0.020	[-0.019, 0.065]	0.277	
block	0.016	0.003	[0.011, 0.022]	< <b>0.001</b>	
stability × condition	-0.211	0.106	[-0.419, -0.003]	<b>0.047</b>	
stability × block	-0.018	0.015	[-0.048, 0.013]	0.252	
condition × block	0.006	0.004	[-0.001, 0.014]	0.080	
stability × condition × block	0.034	0.029	[-0.023, 0.091]	0.242	
<b>hippocampus</b> †					
stability	0.265	0.098	[0.072, 0.458]	<b>0.007</b>	
condition	0.033	0.013	[0.007, 0.059]	<b>0.013</b>	
block	0.017	0.002	[0.013, 0.020]	< <b>0.001</b>	
stability × condition	-0.182	0.192	[-0.558, 0.195]	0.344	
stability × block	-0.004	0.027	[-0.057, 0.050]	0.890	
condition × block	0.006	0.004	[-0.002, 0.013]	0.146	
stability × condition × block	0.101	0.055	[-0.007, 0.208]	0.067	
<b>premotor cortex</b>					
stability	0.349	0.059	[0.226, 0.472]	< <b>0.001</b>	
condition	0.028	0.021	[-0.015, 0.070]	0.191	
block	0.017	0.003	[0.011, 0.023]	< <b>0.001</b>	
stability × condition	-0.102	0.090	[-0.280, 0.075]	0.256	
stability × block	-0.006	0.013	[-0.031, 0.020]	0.660	
condition × block	0.006	0.004	[-0.001, 0.014]	0.080	
stability × condition × block	0.026	0.023	[-0.020, 0.071]	0.265	
<b>LOC</b> †					
stability	0.082	0.045	[-0.006, 0.169]	0.067	
condition	0.033	0.013	[0.007, 0.059]	<b>0.013</b>	
block	0.018	0.002	[0.014, 0.022]	< <b>0.001</b>	
stability × condition	0.033	0.068	[-0.100, 0.167]	0.624	

stability × block	0.004	0.010	[-0.015, 0.023]	0.708
condition × block	0.005	0.004	[-0.002, 0.013]	0.171
stability × condition × block	0.002	0.019	[-0.036, 0.040]	0.919

**Appendix B, Table 5.** Results from Bayesian multilevel regression models examining the relationship between order memory and within-event stability across conditions and blocks. Models include random slopes for all main effects, as well as random intercepts for each participant. Directional posterior probability (*pd*) refers to the proportion of each model term's posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd*> 0.975).

order memory ~ stability × condition × block (Bayesian)					
	Fixed Effect	Estimate	Std. Error	95% CrI	pd
<b>hippocampus</b>					
	stability	0.288	0.103	[0.082, 0.490]	<b>0.997</b>
	condition	0.034	0.021	[-0.009, 0.076]	0.938
	block	0.017	0.003	[0.011, 0.022]	<b>1.000</b>
	stability × condition	-0.157	0.199	[-0.533, 0.222]	0.784
	stability × block	-0.002	0.028	[-0.057, 0.052]	0.527
	condition × block	0.006	0.004	[-0.001, 0.013]	0.940
	stability × condition × block	0.095	0.054	[-0.012, 0.202]	0.959
<b>LOC</b>					
	stability	0.045	0.051	[-0.057, 0.147]	0.811
	condition	0.035	0.021	[-0.008, 0.077]	0.951
	block	0.017	0.003	[0.011, 0.023]	<b>1.000</b>
	stability × condition	0.031	0.081	[-0.127, 0.186]	0.647
	stability × block	0.006	0.011	[-0.016, 0.027]	0.698
	condition × block	0.006	0.004	[-0.002, 0.013]	0.931
	stability × condition × block	0.001	0.020	[-0.037, 0.039]	0.529

*CrI* = credibility interval

*pd* = directional posterior probability

**Appendix B, Table 6.** Results from Frequentist multilevel regression models examining the relationship between within-event neural stability in the hippocampus and template activation in cortex across conditions and blocks. Models include random intercepts for each participant. (For corresponding results from Bayesian regression models including random slopes, see **Appendix B, Table 7.**)

hippocampal stability ~ template activation × condition × block					
	Fixed Effect	Estimate	Std. Error	95% CI	p
<b>mPFC</b>					
template act	0.030	0.009	[0.011, 0.048]	<b>0.002</b>	
condition	0.001	0.004	[-0.006, 0.008]	0.847	
block	0.0005	0.001	[-0.001, 0.002]	0.382	
template act × condition	0.006	0.018	[-0.028, 0.041]	0.716	
template act × block	0.005	0.002	[0.000, 0.009]	<b>0.047</b>	
condition × block	-0.002	0.001	[-0.004, -0.000]	<b>0.029</b>	
template act × condition × block	-0.002	0.005	[-0.011, 0.008]	0.719	
<b>PMC</b>					
template act	0.014	0.009	[-0.004, 0.033]	0.126	
condition	0.002	0.004	[-0.006, 0.009]	0.645	
block	0.0004	0.001	[-0.001, 0.001]	0.482	
template act × condition	0.054	0.017	[0.020, 0.088]	<b>0.002</b>	
template act × block	0.003	0.002	[-0.002, 0.007]	0.244	
condition × block	-0.002	0.001	[-0.004, 0.000]	0.076	
template act × condition × block	0.002	0.005	[-0.007, 0.012]	0.655	
<b>premotor cortex</b>					
template act	0.009	0.010	[-0.011, 0.029]	0.370	
condition	0.002	0.004	[-0.006, 0.009]	0.669	
block	0.0004	0.001	[-0.001, 0.001]	0.503	
template act × condition	0.051	0.017	[0.018, 0.083]	<b>0.002</b>	
template act × block	-0.002	0.002	[-0.007, 0.002]	0.339	
condition × block	-0.001	0.001	[-0.003, 0.001]	0.284	
template act × condition × block	0.000	0.005	[-0.009, 0.010]	0.928	

**Appendix B, Table 7.** Results from Bayesian multilevel regression models examining the relationship between within-event neural stability in the hippocampus and template activation in cortex across conditions and blocks. Models include random slopes for all main effects, as well as random intercepts for each participant. Directional posterior probability (*pd*) refers to the proportion of each model term's posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd* > 0.975).

**hippocampal stability ~ template activation × condition × block (Bayesian)**

	Fixed Effect	Estimate	Std. Error	95% CrI	pd
<b>mPFC</b>					
	template act	0.029	0.013	[0.002, 0.055]	<b>0.981</b>
	condition	0.001	0.004	[-0.007, 0.008]	0.563
	block	0.001	0.001	[-0.001, 0.002]	0.813
	template act × condition	0.006	0.018	[-0.030, 0.042]	0.622
	template act × block	0.004	0.002	[-0.001, 0.009]	0.958
	condition × block	-0.002	0.001	[-0.004, -0.000]	<b>0.977</b>
	template act × condition × block	-0.001	0.005	[-0.011, 0.008]	0.600
<b>PMC</b>					
	template act	0.016	0.012	[-0.009, 0.041]	0.901
	condition	0.001	0.004	[-0.008, 0.010]	0.617
	block	0.000	0.001	[-0.001, 0.002]	0.746
	template act × condition	0.051	0.019	[0.016, 0.088]	<b>0.998</b>
	template act × block	0.003	0.002	[-0.002, 0.008]	0.891
	condition × block	-0.002	0.001	[-0.004, 0.000]	0.966
	template act × condition × block	0.002	0.005	[-0.008, 0.011]	0.621
<b>premotor cortex</b>					
	template act	0.007	0.011	[-0.015, 0.029]	0.736
	condition	0.002	0.005	[-0.008, 0.011]	0.646
	block	0.000	0.001	[-0.001, 0.002]	0.700
	template act × condition	0.061	0.019	[0.024, 0.100]	<b>0.999</b>
	template act × block	-0.003	0.003	[-0.008, 0.002]	0.873
	condition × block	-0.001	0.001	[-0.003, 0.001]	0.810
	template act × condition × block	0.001	0.005	[-0.008, 0.010]	0.579

*CrI* = credibility interval

*pd* = directional posterior probability

**Appendix B, Table 8.** Results from Frequentist multilevel regression model examining how within-event neural stability in the lateral occipital cortex (LOC) is affected by template activation, condition, and block. Models include random intercepts for each participant. (For corresponding results from Bayesian regression models including random slopes, see **Appendix B, Table 9.**)

LOC stability ~ template activation × condition × block					
	Fixed Effect	Estimate	Std. Error	95% CI	p
<b>mPFC</b>					
template act	0.045	0.021	[0.004, 0.085]	<b>0.033</b>	
condition	0.008	0.008	[-0.008, 0.023]	0.329	
block	-0.010	0.001	[-0.012, -0.008]	< <b>0.001</b>	
template act × condition	0.032	0.038	[-0.043, 0.107]	0.398	
template act × block	0.010	0.005	[-0.000, 0.020]	0.061	
condition × block	-0.001	0.002	[-0.005, 0.004]	0.815	
template act × condition × block	-0.000	0.010	[-0.020, 0.020]	0.983	
<b>PMC</b>					
template act	0.044	0.021	[0.003, 0.085]	<b>0.037</b>	
condition	0.008	0.008	[-0.008, 0.023]	0.333	
block	-0.010	0.001	[-0.012, -0.008]	< <b>0.001</b>	
template act × condition	0.092	0.037	[0.019, 0.165]	<b>0.013</b>	
template act × block	0.008	0.005	[-0.002, 0.018]	0.096	
condition × block	-0.000	0.002	[-0.004, 0.004]	0.987	
template act × condition × block	0.011	0.010	[-0.008, 0.031]	0.259	
<b>premotor cortex</b>					
template act	0.070	0.024	[0.023, 0.117]	<b>0.003</b>	
condition	0.003	0.008	[-0.013, 0.019]	0.730	
block	-0.009	0.001	[-0.012, -0.007]	< <b>0.001</b>	
template act × condition	0.055	0.036	[-0.015, 0.125]	0.126	
template act × block	-0.004	0.005	[-0.014, 0.006]	0.425	
condition × block	0.001	0.002	[-0.003, 0.006]	0.595	
template act × condition × block	0.008	0.010	[-0.011, 0.027]	0.413	

**Appendix B, Table 9.** Results from Bayesian multilevel regression model examining how within-event neural stability in the lateral occipital cortex (LOC) is affected by template activation, condition, and block. Models include random slopes for all main effects, as well as random intercepts for each participant. Directional posterior probability (*pd*) refers to the proportion of each model term's posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd* > 0.975).

LOC stability ~ template activation × condition × block (Bayesian)

	Fixed Effect	Estimate	Std. Error	95% CrI	pd
<b>mPFC</b>					
	template act	0.047	0.027	[-0.006, 0.102]	0.961
	condition	0.007	0.015	[-0.022, 0.036]	0.688
	block	-0.010	0.002	[-0.014, -0.006]	<b>1.000</b>
	template act × condition	0.021	0.042	[-0.061, 0.103]	0.686
	template act × block	0.005	0.005	[-0.006, 0.016]	0.821
	condition × block	-0.000	0.002	[-0.005, 0.004]	0.559
	template act × condition × block	0.007	0.010	[-0.013, 0.028]	0.747
<b>PMC</b>					
	template act	0.053	0.022	[0.008, 0.098]	<b>0.990</b>
	condition	0.006	0.015	[-0.024, 0.036]	0.665
	block	-0.010	0.002	[-0.014, -0.006]	<b>1.000</b>
	template act × condition	0.096	0.043	[0.011, 0.178]	<b>0.987</b>
	template act × block	0.008	0.006	[-0.003, 0.019]	0.913
	condition × block	0.000	0.002	[-0.004, 0.004]	0.540
	template act × condition × block	0.012	0.010	[-0.008, 0.032]	0.882
<b>premotor cortex</b>					
	template act	0.123	0.029	[0.066, 0.181]	<b>1.000</b>
	condition	-0.004	0.017	[-0.039, 0.030]	0.604
	block	-0.009	0.002	[-0.013, -0.005]	> <b>0.999</b>
	template act × condition	0.092	0.048	[-0.001, 0.187]	0.974
	template act × block	-0.003	0.006	[-0.015, 0.008]	0.711
	condition × block	0.002	0.002	[-0.003, 0.006]	0.803
	template act × condition × block	0.008	0.010	[-0.012, 0.027]	0.787

*CrI* = credibility interval

*pd* = directional posterior probability

**Appendix B, Table 10.** Results from Frequentist multilevel regression models examining the relationship between within-event neural stability and template activation *within the same ROI* across conditions and blocks. Models include random intercepts for each participant. (For analogous results from Bayesian regression models including random slopes, see **Appendix B, Table 11.**)

stability ~ template activation × condition × block (stability & template activation in the same region)					
	Fixed Effect	Estimate	Std. Error	95% CI	p
<b>mPFC</b>					
template act	0.091	0.018	[0.055, 0.127]	< <b>0.001</b>	
condition	0.003	0.007	[-0.010, 0.016]	0.671	
block	-0.001	0.001	[-0.003, 0.001]	0.545	
template act × condition	-0.021	0.033	[-0.086, 0.044]	0.528	
template act × block	0.012	0.004	[0.003, 0.021]	<b>0.007</b>	
condition × block	-0.001	0.002	[-0.005, 0.003]	0.500	
template act × condition × block	-0.009	0.009	[-0.026, 0.009]	0.321	
<b>PMC</b>					
template act	0.082	0.016	[0.050, 0.115]	< <b>0.001</b>	
condition	0.017	0.006	[0.005, 0.029]	<b>0.005</b>	
block	0.001	0.001	[-0.001, 0.003]	0.352	
template act × condition	0.065	0.029	[0.007, 0.122]	<b>0.027</b>	
template act × block	0.013	0.004	[0.005, 0.021]	<b>0.001</b>	
condition × block	-0.001	0.002	[-0.005, 0.002]	0.456	
template act × condition × block	0.012	0.008	[-0.004, 0.028]	0.137	
<b>premotor cortex</b>					
template act	0.074	0.021	[0.033, 0.114]	< <b>0.001</b>	
condition	0.009	0.007	[-0.005, 0.022]	0.229	
block	0.001	0.001	[-0.001, 0.003]	0.410	
template act × condition	0.028	0.031	[-0.033, 0.088]	0.371	
template act × block	0.012	0.004	[0.004, 0.021]	<b>0.004</b>	
condition × block	-0.002	0.002	[-0.006, 0.002]	0.278	
template act × condition × block	0.007	0.008	[-0.009, 0.024]	0.396	

**Appendix B, Table 11.** Results from Bayesian multilevel regression models examining the relationship between within-event neural stability and template activation *within the same ROI* across conditions and blocks. Models include random slopes for all main effects, as well as random intercepts for each participant. Directional posterior probability (*pd*) refers to the proportion of each model term's posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd* > 0.975).

**stability ~ template activation × condition × block**  
 (stability & template activation in the same region)

	Fixed Effect	Estimate	Std. Error	95% CrI	<i>pd</i>
<b><i>mPFC</i></b>					
	template act	0.094	0.032	[0.030, 0.158]	<b>0.997</b>
	condition	0.004	0.008	[-0.013, 0.020]	0.677
	block	-0.001	0.002	[-0.004, 0.002]	0.313
	template act × condition	-0.015	0.036	[-0.084, 0.056]	0.656
	template act × block	0.011	0.005	[0.001, 0.020]	<b>0.987</b>
	condition × block	-0.001	0.002	[-0.005, 0.003]	0.699
	template act × condition × block	-0.005	0.009	[-0.022, 0.013]	0.696
<b><i>PMC</i></b>					
	template act	0.089	0.034	[0.021, 0.158]	<b>0.995</b>
	condition	0.015	0.008	[-0.000, 0.030]	0.972
	block	0.001	0.001	[-0.001, 0.004]	0.818
	template act × condition	0.053	0.033	[-0.012, 0.117]	0.946
	template act × block	0.011	0.004	[0.002, 0.020]	<b>0.994</b>
	condition × block	-0.001	0.002	[-0.005, 0.002]	0.776
	template act × condition × block	0.005	0.008	[-0.011, 0.021]	0.738
<b><i>premotor cortex</i></b>					
	template act	0.086	0.038	[0.010, 0.165]	<b>0.987</b>
	condition	-0.001	0.014	[-0.031, 0.028]	0.529
	block	0.001	0.002	[-0.003, 0.004]	0.653
	template act × condition	0.115	0.041	[0.034, 0.197]	<b>0.998</b>
	template act × block	0.008	0.005	[-0.002, 0.018]	0.934
	condition × block	-0.000	0.002	[-0.004, 0.003]	0.595
	template act × condition × block	0.009	0.009	[-0.007, 0.026]	0.868

## Appendix C: Chapter 3 Supplement

### C.1 Supplemental Methods & Results

#### C.1.1 GPT-4 prompt for detail segmentation

Your goal is to parse, categorize, and score the valence of the following description. First, parse the description into discrete meaningful chunks.

A chunk is defined as a unique occurrence, observation, or thought, typically expressed as a grammatical unit such as a verb phrase, clause, adjective, proper noun, or separated by a conjunction such as for, and, nor, but, or, yet, so. When parsing, if a sentence or phrase appears to belong to multiple categories, split it into the respective distinct chunks. For example, if “Today I went to the endocrinologist.” can be categorized as both “Event” and “Time”, split it into “Today” as “Time” and “I went to the endocrinologist” as “Event.”

Then, categorize each chunk of the description as Event, Time, Place, Perceptual, Thought/Emotion, Semantic, Extraneous, or Reflection. Each chunk can be assigned to only one category:

- Event: Happenings, individuals present, actions or reactions of self and others, topics of conversation.
- Time: Year, season, month, day of week, time of day, statements about relative timing within an event.
- Place: Localization of an event including the city, street, building, room, part of room.
- Perceptual: Auditory, olfactory, tactile, taste, visual and visual details, body position, duration, weather.
- Thought/Emotion: Emotional state, thoughts, implications of self and others.
- Semantic: General knowledge or facts and pieces of information that are not directly related to the main experience being described.
- Extraneous: Tangential events to the main memory recalled.
- Reflection: Meta-cognitive statements, editorializing, speculation.

After categorizing, score the valence of each chunk as (positive), (negative), or (neutral). Valence captures the emotional tone of the chunk, indicating whether it is positive, negative, or neutral in nature. Items are only positive or negative in nature if they have descriptive adjectives in them.

Please format your response in the following manner:

For chunks:

- Chunk of text
- Another chunk of text,
- ...

For categories:

- Category
- Another category,
- ...

For categories:

- Valence
- Valence,
- ...

### C.1.2 Supplemental Bayesian regression models

In some cases, due to the complexity of the structure of multilevel regression models (and due to our relatively modest sample size), we were unable to estimate models with the full random effects structure (i.e., random intercepts and random slopes for each predictor, nested within participants) using traditional Frequentist methods. We addressed this issue first by simplifying the random effects included in each model (see *Methods*). However, when significant effects were observed in these simplified models, we sought to further verify their validity by using Bayesian estimation methods to fit regression models that did include random slopes. All Bayesian regression models were computed using R's *brms* (ver. 2.21.2) package. Each model used uninformative priors for all parameters, and was fit using 4 Markov chains with 30,000 iterations per chain. Although the Bayesian approach does not use *p*-values to draw a binary distinction between significant and non-significant results, here we assess the statistical reliability of each effect by examining whether its 95% credibility interval (*CrI*) contains zero. A 95% *CrI* that excludes zero can be interpreted similarly to a situation in which  $p < .05$ . We also report directional posterior probabilities (*pd*) for each model term, which describe the proportion of each effect's posterior distribution that falls above (for positive effects) or below (for negative effects) zero. In general, these Bayesian models reproduced what was found with simplified Frequentist models, building confidence that our findings hold true even when fully accounting for variability across subjects.

### C.1.3 Characterizing effects across participant groups

Participants were recruited in two waves of data collection, referred to as Group A ( $N = 15$ ) and Group B ( $N = 26$ ). Each group completed largely identical experimental tasks, with some exceptions — most notably, that participants in Group B received more extensive training about how to identify and describe events during daily diaries (see *Methods* for details). To maximize statistical power, all analyses in Chapter 3 were conducted on data collapsed across groups. However, for the sake of transparency, here we describe relevant differences in behavior between the two groups, and assess whether effects of interest replicate when looking at each group separately. For analyses based on t-tests, we report results

from two-sample t-tests and do not assume equal variances (i.e., Welch's t-tests). For analyses based on regression models, we use a Frequentist multilevel modeling approach, where each model included a random intercept for each participant. Because running separate models for each group substantially reduced the amount of data included within each model, random slopes could not be estimated reliably (with Frequentist or Bayesian methods) and thus were not included.

**Daily diary behavior.** Due to the additional guidance we provided to participants in Group B on reporting events in a clear and detailed manner, the descriptions they generated tended to be longer than those generated by Group A — both in terms of the number of words ( $t(29.4) = -8.286$ ,  $CI = [-85.732, -51.802]$ ,  $p < .001$ ,  $d = 2.090$ ) and the number of episodic details ( $t(36.8) = -9.159$ ,  $CI = [-6.297, -4.016]$ ,  $p < .001$ ,  $d = 2.454$ ) per event. However, the quantity of novel experiences recorded by each group during daily diaries was generally comparable. Although participants in Group B reported a larger proportion of “new” events ( $t(39.0) = -2.889$ ,  $CI = [-0.173, -0.030]$ ,  $p = .006$ ,  $d = 0.814$ ), the two groups were similar in the proportion of days in which a novel location was visited ( $t(28.8) = -0.176$ ,  $CI = [-0.138, 0.116]$ ,  $p = .861$ ,  $d = 0.057$ ), the proportion of days in which a novel social interaction occurred ( $t(31.5) = -0.625$ ,  $CI = [-0.203, 0.108]$ ,  $p = .537$ ,  $d = 0.198$ ), and the proportion of days that were described as “notable” ( $t(36.2) = -1.445$ ,  $CI = [-0.265, 0.045]$ ,  $p = .157$ ,  $d = 0.434$ ). The two groups also did not differ in average ratings of day typicality ( $t(39.0) = -1.354$ ,  $CI = [-0.455, 0.090]$ ,  $p = .184$ ,  $d = 0.383$ ).

**Memory test behavior.** Event memory was measured via two metrics: subjective memory vividness scores and the number of episodic details recalled per event. Although vividness scores were similar across groups ( $t(32.7) = -0.861$ ,  $CI = [-0.515, 0.209]$ ,  $p = .396$ ,  $d = 0.269$ ), Group B participants tended to recall more episodic details ( $t(31.5) = -2.165$ ,  $CI = [-2.733, -0.083]$ ,  $p = .038$ ,  $d = 0.685$ ) — again, likely due to methodological changes we implemented for this group (see *Methods*). Interestingly, however, a mixed ANOVA revealed that although Group B participants included more episodic details during both when describing events during the daily diary period and when recalling them during the recall test (main effect of group:  $F(1, 39) = 31.592$ ,  $p < .001$ ), the magnitude of this difference was significantly smaller at recall relative to during the diary period (group-by-timepoint interaction:  $F(1, 39)$

$= 25.929, p < .001$ ). When examining how event memory changed based on the delay between each event and the memory test, effects in each group were similar as what we observed across all participants. That is, memory vividness decreased with longer delays in Group A ( $b = -0.032, SE = 0.013, CI = [-0.058, -0.006], p = .016$ ) and Group B ( $b = -0.020, SE = 0.009, CI = [-0.039, -0.002], p = .030$ ). In contrast, the number of episodic details recalled did not vary based on delay in either group (Group A:  $b = -0.051, SE = 0.039, CI = [-0.127, 0.026], p = .192$ ; Group B:  $b = -0.020, SE = 0.029, CI = [-0.078, 0.038], p = .491$ ). Also consistent with what was observed when collapsing across groups, the number of details recalled per event was positively associated with both subjective memory vividness ratings (Group A:  $b = 0.264, SE = 0.036, CI = [0.193, 0.336], p < .001$ ; Group B:  $b = 0.201, SE = 0.025, CI = [0.153, 0.250], p < .001$ ) and the number of details reported during the diary period (Group A:  $b = 0.379, SE = 0.115, CI = [0.152, 0.606], p = .001$ ; Group B:  $b = 0.129, SE = 0.044, CI = [0.042, 0.216], p = .004$ ).

We then examined group-related differences in temporal memory behavior, i.e., temporal order accuracy and subjective temporal distance ratings. Temporal order accuracy did not significantly differ between Group A and Group B ( $t(27.9) = 0.574, CI = [-0.071, 0.126], p = .571, d = 0.189$ ), and both groups showed a strong positive association between order accuracy and the confidence ratings provided for each order judgment (Group A:  $b = 0.504, SE = 0.128, CI = [0.254, 0.755], p < .001$ ; Group B:  $b = 0.230, SE = 0.086, CI = [0.121, 0.458], p < .001$ ). Next, we assessed whether order accuracy increased based on the vividness with which participants reported remembering each individual event in a given pair. In Group A, order accuracy did not reliably increase as a function of the average memory vividness of constituent events ( $b = 0.140, SE = 0.156, CI = [-0.165, 0.445], p = .368$ ). However, this effect did reach significance in Group B ( $b = 0.256, SE = 0.112, CI = [0.037, 0.475], p = .022$ ). Finally, as with the effect observed across both groups, temporal order memory did not vary based on the delay between when each event occurred and when memory was tested (Group A:  $b = 0.044, SE = 0.039, CI = [-0.034, 0.121], p = .268$ ; Group B:  $b = -0.002, SE = 0.032, CI = [-0.066, 0.061], p = .942$ )

We then turned to ratings of subjective temporal distance (i.e., how far apart in time two events are felt to be). On average, participants in Group A produced more compressed ratings of temporal

distance than those in Group B ( $t(27.0) = -3.791$ ,  $CI = [-0.696, -0.549]$ ,  $p < .001$ ,  $d = 1.263$ ). Finally, although both groups reported numerically higher distance ratings for event pairs separated by 3-4 days versus 1-2 days, this contrast only reached significance in Group A ( $t(14) = -2.249$ ,  $CI = [-0.378, -0.009]$ ,  $p = .041$ ,  $d = 0.465$ ) and not Group B ( $t(25) = -1.405$ ,  $CI = [-0.371, 0.070]$ ,  $p = .17$ ,  $d = 0.347$ ).

***Relationships between novelty and event memory performance.*** In Chapter 3, we reported that both subjective memory vividness scores and the number of episodic details recalled varied as a function of the regularity of each event, such that “new” events were associated with stronger memory recall, followed by “periodic” events, with “routine” events remembered in the least amount of detail. We observed this same pattern of effects when looking at memory vividness scores in both Group A (new vs. periodic:  $b = 1.016$ ,  $SE = 0.143$ ,  $CI = [0.679, 1.352]$ ,  $p_{tukey} < .001$ ; new vs. routine:  $b = 2.032$ ,  $SE = 0.168$ ,  $CI = [1.638, 2.426]$ ,  $p_{tukey} < .001$ ; periodic vs. routine:  $b = 1.016$ ,  $SE = 0.124$ ,  $CI = [0.724, 1.308]$ ,  $p_{tukey} = .004$ ) and Group B (new vs. periodic:  $b = 0.696$ ,  $SE = 0.090$ ,  $CI = [0.485, 0.907]$ ,  $p_{tukey} < .001$ ; new vs. routine:  $b = 1.276$ ,  $SE = 0.140$ ,  $CI = [0.948, 1.604]$ ,  $p_{tukey} < .001$ ; periodic vs. routine:  $b = 0.508$ ,  $SE = 0.124$ ,  $CI = [0.298, 0.872]$ ,  $p_{tukey} < .001$ ). Although the number of episodic details recalled about each event also increased as a function of the event’s novelty in each participant group, effects were weaker overall. In Group A, both “new” events and “periodic” events were recalled with significantly more details than “routine” events (new vs. routine:  $b = 2.245$ ,  $SE = 0.529$ ,  $CI = [0.993, 3.497]$ ,  $p_{tukey} < .001$ ; periodic vs. routine:  $b = 1.343$ ,  $SE = 0.386$ ,  $CI = [0.431, 2.256]$ ,  $p_{tukey} = .002$ ), but the contrast between “new” and “periodic” events was only marginal ( $b = 0.901$ ,  $SE = 0.464$ ,  $CI = [-0.196, 1.998]$ ,  $p_{tukey} = .130$ ). In Group B, participants recalled “new” events in marginally more detail than “routine” events ( $b = 0.985$ ,  $SE = 0.475$ ,  $CI = [-0.133, 2.103]$ ,  $p_{tukey} = .097$ ), with no significant difference in memory for “new” versus “periodic” events ( $b = 0.576$ ,  $SE = 0.294$ ,  $CI = [-0.116, 1.269]$ ,  $p_{tukey} = .124$ ), or between “periodic” and “routine” events ( $b = 0.408$ ,  $SE = 0.431$ ,  $CI = [-0.606, 1.422]$ ,  $p_{tukey} = .610$ ).

Next, we examined how memory vividness ratings were associated with the full set of novelty variables captured as part of participants’ daily diaries. These novelty variables included *event regularity*, *day typicality*, *novel locations visited*, and *novel social interactions*, and *day notability* (see *Methods* for

details). Consistent with our main findings, in both Group A and Group B, memory vividness scores were significantly associated with event regularity, such that “new” events were remembered more vividly than “periodic” events (Group A:  $b = 0.865$ ,  $SE = 0.145$ ,  $CI = [0.524, 1.205]$ ,  $p_{tukey} < .001$ ; Group B:  $b = 0.624$ ,  $SE = 0.091$ ,  $CI = [0.411, 0.838]$ ,  $p_{tukey} < .001$ ) and “routine” events (Group A:  $b = 1.836$ ,  $SE = 0.172$ ,  $CI = [1.433, 2.240]$ ,  $p_{tukey} < .001$ ; Group B:  $b = 1.178$ ,  $SE = 0.141$ ,  $CI = [0.847, 1.509]$ ,  $p_{tukey} < .001$ ), and “periodic” events were remembered more vividly than “routine” events (Group A:  $b = 0.971$ ,  $SE = 0.124$ ,  $CI = [0.680, 1.263]$ ,  $p_{tukey} < .001$ ; Group B:  $b = 0.554$ ,  $SE = 0.124$ ,  $CI = [0.264, 0.844]$ ,  $p_{tukey} < .001$ ). Vividness was also higher in both groups when a novel location had been visited that day (Group A:  $b = 0.351$ ,  $SE = 0.117$ ,  $CI = [0.121, 0.581]$ ,  $p = .003$ ; Group B:  $b = 0.230$ ,  $SE = 0.091$ ,  $CI = [0.051, 0.409]$ ,  $p = .012$ ), and when a day was considered “notable” (Group A:  $b = 0.393$ ,  $SE = 0.124$ ,  $CI = [0.150, 0.636]$ ,  $p = .002$ ; Group B:  $b = 0.190$ ,  $SE = 0.090$ ,  $CI = [0.011, 0.368]$ ,  $p = .037$ ). No significant associations were found between memory vividness and novel social interactions occurring on that day, or between vividness and ratings of day typicality (all  $p > .3$ ).

The same set of novelty variables was then used to predict the number of episodic details remembered per event. In Group A, similar to results across both groups, participants recalled more episodic details for “new” relative to “routine” events ( $b = 2.014$ ,  $SE = 0.678$ ,  $CI = [0.678, 3.351]$ ,  $p_{tukey} = .001$ ) and for “routine” relative to “periodic” events ( $b = 1.336$ ,  $SE = 0.404$ ,  $CI = [0.380, 2.292]$ ,  $p_{tukey} = .003$ ), but not for “new” relative to “periodic” events ( $b = 0.678$ ,  $SE = 0.488$ ,  $CI = [-0.475, 1.832]$ ,  $p_{tukey} = .348$ ). However, no other novelty variables showed a significant effect on detail counts (all  $p > .1$ ). In Group B, there were no significant differences in the number of details recalled based on event regularity (all  $p_{tukey} > .2$ ), nor were more details recalled for days described as “notable”, days in which a novel location was visited, or days with lower typicality ratings (all  $p > .4$ ). In contrast, detail counts in Group B were significantly higher for days in which a novel social interaction occurred ( $b = 0.667$ ,  $SE = 0.285$ ,  $CI = [0.107, 1.227]$ ,  $p = 0.020$ ).

***Relationships between novelty and temporal memory performance.*** In the final section of results, we examined whether temporal memory for a given pair of events was affected by what degree of

novelty participants were exposed to throughout the days *between* those events. Specifically, we constructed a set of models in which temporal order accuracy and subjective temporal distance were predicted by the number of new experiences during intervening days as well as the full set of our day-level novelty metrics (i.e., the mean typicality of intervening days, the number of intervening days where novel locations were visited, the number of intervening days where a novel social interaction occurred, and the number of days considered “notable”). Both models included a regressor for the objective distance (in days) between the events in each pair. Temporal accuracy showed no association with any of these examined variables, both in Group A (all  $p > .08$ ) and Group B (all  $p > .3$ ). In a model where subjective temporal distance ratings were predicted by all novelty variables, we saw in Group A that reporting more novel experiences in the interval between two events was associated with shorter distance judgments ( $b = -0.205$ ,  $SE = 0.081$ ,  $CI = [-0.364, -0.047]$ ,  $p = .012$ ), with no other significant effects (all  $p > .1$ ). In Group B, there were no significant associations between distance judgments and these novelty variables (all  $p > .2$ ).

#### C.1.4 Number of details recalled by event regularity: controlling for diary entry details

As noted in Chapter 3, we found that the number of episodic details recalled per event was significantly modulated by event regularity, such that more regular (i.e., “routine” or “periodic”) events were generally recalled in less detail relative to more novel experiences. However, we also found that detail counts during recall were positively related to the amount of episodic information recorded by participants *during daily diary* (i.e., on the day that the event occurred) — raising the possibility that the effects we see here are driven by differences in how participants describe each type of the event during the diary period. To address this, we ran additional models in which the number of details recalled per event was predicted by event regularity while controlling for the number of details reported originally during the diary period. We used a Bayesian regression model (with random intercepts and slopes) for this purpose, as a comparable Frequentist model could not be reliably estimated. (See *Methods* for details about how we assessed the statistical robustness of each Bayesian model term.) Importantly, this model replicated the effects observed in our main analysis: namely, that participants recalled more details for

“new” events compared to both “periodic” events ( $b = -0.545$ ,  $SE = 0.270$ ,  $CrI = [-1.075, -0.013]$ ,  $pd = .978$ ) and “routine” events ( $b = -1.461$ ,  $SE = 0.365$ ,  $CrI = [-2.173, -0.740]$ ,  $pd = 1.000$ ). As expected, there was also a strong positive association between the number of details reported during recall and during the original diary period ( $b = 0.168$ ,  $SE = 0.053$ ,  $CrI = [0.066, 0.273]$ ,  $pd = 0.999$ ).

### C.1.5 Effects of novelty on memory vividness for non-novel events from adjacent days

Our main results indicate that engaging in a novel experience enhances memory vividness for non-novel (i.e., “periodic” and “routine”) events that were reported on the same day. To explore the temporal extent of this effect, we conducted an exploratory analysis asking whether vividness scores were also higher for non-novel experiences reported one day before or after a new event — excluding any events that came from the same day as a new event. However, there was no difference in memory vividness scores for routine and periodic events that occurred on adjacent days to new events ( $b = -0.154$ ,  $SE = 0.115$ ,  $CI = [-0.388, 0.080]$ ,  $p = 0.190$ ).

## C.2 Supplemental Tables

**Appendix C, Table 1.** Pre-diary and post-diary questionnaire sections.

questionnaire section	overview of questionnaire items
demographics <i>(only included in pre-diary questionnaire)</i>	<ul style="list-style-type: none"> <li>• age</li> <li>• gender</li> <li>• ethnicity</li> <li>• race</li> <li>• subjective social status (using MacArthur Scale of Subjective Social Status)</li> <li>• highest level of education</li> </ul>
effects of COVID-19	<ul style="list-style-type: none"> <li>• concern about self/close others becoming infected with COVID-19</li> <li>• whether self/close others belong to an “at risk” group</li> <li>• whether self/close others work as a first responder</li> <li>• predictions about whether the COVID-19 situation will improve/worsen in the next 1/2/6/12 months</li> <li>• current stress due to consequences of COVID-19 (e.g., lack of in-person contact, financial issues, mental &amp; physical health issues, lack of sufficient government action, etc.)</li> <li>• predictions about future stress due to consequences of COVID-19</li> <li>• perceived changes in emotional state relative to before the pandemic</li> <li>• perceived changes in behavior relative to before the pandemic</li> <li>• consumption of COVID-19 news</li> </ul>
mental health status	<ul style="list-style-type: none"> <li>• history of counseling/treatment for mental health difficulties</li> <li>• history of mental health diagnoses</li> </ul>
<i>Note that all questions in this section asked participants to respond to yes / no questions. Detailed information about mental health history was not collected.</i>	
typical daily behavior	<ul style="list-style-type: none"> <li>• list of events/activities experienced on most days</li> <li>• typical frequency of leaving the house/visiting new locations</li> <li>• typical frequency of social interactions</li> <li>• typical frequency of news consumption</li> <li>• typical experience of time (e.g., how quickly the day passes)</li> <li>• perceived memorability of recent experience</li> <li>• perceived impact of COVID-19 on daily behavior</li> </ul>
State-Trait Anxiety Inventory, Trait Version (STAI-T)	self-report questionnaire measuring an individual’s general tendency to experience anxiety
Positive and Negative Affect Schedule (PANAS)	self-report questionnaire measuring an individual’s current emotional state
Intolerance of Uncertainty Scale (IUS)	self-report questionnaire measuring how an individual feels about and reacts to the uncertainties of life
Depression Anxiety Stress Scale (DASS-21)	self-report questionnaire measuring the extent to which an individual experienced depression, anxiety, and stress over the last week

**Appendix C, Table 2.** “Day-level” variables collected as part of each daily diary. By day-level, we refer to variables that were measured once per day (and describe the participant’s behavior throughout that day).

<b>Q#</b>	<b>diary item</b>	<b>question text</b>	<b>response format</b>
1	current mood	Please rate each item below based on how you feel right now: <ul style="list-style-type: none"> <li>• Positive</li> <li>• Calm</li> <li>• Negative</li> <li>• Afraid</li> <li>• Excited</li> <li>• Happy</li> <li>• Sad</li> </ul>	<b>1-5 scale</b> (for each emotion) 1 = not at all 5 = a great deal
2	current location	Where are you currently?	<b>multiple choice</b> <ul style="list-style-type: none"> <li>• in my home town/city/environment</li> <li>• out of town (with family)</li> <li>• out of town (with friends)</li> <li>• out of town (for work)</li> <li>• other [fill in the blank]</li> </ul>
3	similarity to yesterday	How similar does today feel to yesterday?	<b>1-5 scale</b> 1 = not at all 5 = a great deal
4	similarity to typical day*	How similar does today feel to the "typical day" you described in the initial questionnaire for this study?	<b>1-5 scale</b> 1 = not at all 5 = a great deal
5	subjective duration	On a scale of 1-5, how would you describe how long today felt?	<b>1-5 scale</b> 1 = “flew by” 5 = “dragged on”
6	number of events	<i>We can describe our days as a series of numeric input events, such that each event is marked by a clear beginning, end, goal, activity, and/or spatial or social context. For example, you may describe "eating breakfast" or "taking my dog for a walk" as two individual events.</i>	
		How many different “events” has your day consisted of?	
7	locations outside home	Did you go anywhere (outside of your home) today?	<b>yes / no</b>
8	number of locations visited (if yes to Q7)	How many different places did you go?	<b>numeric input</b>
9	type of locations visited*	Where did you go?	<b>check all that apply</b> <ul style="list-style-type: none"> <li>• familiar outdoor location</li> <li>• novel outdoor location</li> <li>• familiar commercial location</li> <li>• novel commercial location</li> <li>• familiar residential location</li> <li>• novel residential location</li> </ul>

10	exercise	Did you exercise today?	<b>yes / no</b>
11	step tracker	Do you have a step tracker that you can refer to?	<b>yes / no</b>
12	steps	How many steps have you taken today? <b>numeric input</b>	 <i>If you have a step tracker on your phone, please report that number. If you do not, please approximate, given that 1 mile of walking = ~2000 steps.</i>
13	hours spent alone	How many hours did you spend doing activities alone?	<b>numeric input</b>
14	social interactions	Did you interact with anyone (outside of those you live with) today?	<b>yes / no</b>
15	social interactions: <i>who</i> * <i>(if yes to Q14)</i>	Who did you interact with?	<b>check all that apply</b> <ul style="list-style-type: none"> <li>• spouse/partner</li> <li>• friend(s)</li> <li>• family member(s)</li> <li>• acquaintance(s)</li> <li>• stranger(s)</li> <li>• colleague(s) / coworker(s)</li> </ul>
16	social interactions: <i>regularity</i> *	How often do you interact with this person/these people?	<b>multiple choice</b> <ul style="list-style-type: none"> <li>• this was our first interaction</li> <li>• occasionally</li> <li>• frequently</li> </ul>
17	social interactions: <i>type</i>	Was this interaction virtual or in-person?	<b>multiple choice</b> <ul style="list-style-type: none"> <li>• virtual</li> <li>• in-person</li> <li>• both</li> </ul>
18	social interactions: <i>duration</i>	How long did the interaction last? (in hours)	<b>numeric input</b>
19	more social interactions	Would you like to report another social interaction?	<b>yes / no</b> <i>if yes:</i> repeat Q15-17 (for up to 10 social interactions)
20	day notability*	Did anything notable, out of the ordinary, very important, and/or unexpected happen to you today?	<b>yes / no</b>
21	COVID-19: <i>impact</i>	How did COVID-19 affect your activities today?	<b>check all that apply</b> <ul style="list-style-type: none"> <li>• COVID-19 crossed my mind but did not impact my behavior</li> <li>• COVID-19 prevented me from going somewhere</li> <li>• COVID-19 prevented me from interacting with someone COVID-19 prevented me from doing an activity</li> <li>• I felt worried about being infected with COVID-19</li> <li>• I did not think about COVID-19</li> </ul>

22	COVID-19: <i>worry</i>	How worried about COVID-19 did you <b>1-5 scale</b> feel? 1 = not at all 5 = a great deal
----	------------------------	--

*All questions below were asked after participants reported specific events (see Appendix C, Table 3).*

23	news	Did you read/hear about any piece(s) of <b>yes / no</b> news today?
24	news: <i>story details</i>	Briefly list all news stories you read/heard about (one per line). <b>text box</b>
25	news: <i>engagement</i>	How closely did you follow the news? <b>1-5 scale</b> 1 = not at all 5 = a great deal
26	news: <i>sharing</i>	Did you talk with others about the news today? <b>multiple choice</b> • Yes, a lot. • Yes, a little. • No.

**Appendix C, Table 3.** “Event-level” variables collected as part of each daily diary. By event-level, we refer to variables that were measured once per event (with 3 events measured per diary day).

Q#	diary item	question text	response format
0	event instructions	<i>Now we'll ask you to identify and answer some questions about 3 events/activities that have happened today. An activity is broadly construed as a thing you have done (e.g., listening to music, working on a project, talking on the phone, etc.), either on your own or with others.</i>	
1	event description**	Please describe an event/activity from today. Be as detailed as possible (2-4 sentences).	<b>text box</b>
2	event title**	Please give this event/activity a unique title. Your title should be brief (< 10 words), but distinctive enough to remind you of this specific experience on this particular day.	<b>text box</b>
3	event subjective duration	On a scale of 1-5, how long did this event/activity feel?	<b>1-5 scale</b> 1 = “flew by” 5 = “dragged on”
4	event objective duration	What was the actual duration of this event/activity (in hours)?	<b>numeric input</b>
5	event time of day	When did this event/activity take place?	<b>multiple choice</b> <ul style="list-style-type: none"><li>• morning</li><li>• afternoon</li><li>• evening</li><li>• late night/early morning</li></ul>
6	event social companion(s): <i>category</i>	Who were you with when this event/activity took place?	<b>check all that apply</b> <ul style="list-style-type: none"><li>• alone</li><li>• spouse/partner</li><li>• family member(s)</li><li>• friend(s)</li><li>• acquaintance(s)</li><li>• stranger(s)</li><li>• colleague(s) / coworker(s)</li></ul>
7	event social companion(s): <i>specific</i>	Whom were you with when this event/activity took place, specifically?  <i>Please list first names only with each name separated by a comma. If you were only with strangers please write N/A.</i>	<b>text box</b>
8	event social companion(s): <i>regularity*</i>	How often do you interact with this person/these people?	<b>multiple choice</b> <ul style="list-style-type: none"><li>• this was our first interaction</li><li>• occasionally</li><li>• frequently</li></ul>
9	event social companion(s): <i>interaction type</i>	Did you interact with this person/these people virtually or in-person?	<b>multiple choice</b> <ul style="list-style-type: none"><li>• virtually</li><li>• in-person</li><li>• both</li></ul>

10	event location: <i>category</i> *	In what kind of location did this event/activity take place?	<b>check all that apply</b>
			<ul style="list-style-type: none"> <li>• familiar outdoor location</li> <li>• novel outdoor location</li> <li>• familiar commercial location</li> <li>• novel commercial location</li> <li>• familiar residential location</li> <li>• novel residential location</li> </ul>
11	event location: <i>specific</i>	Where did this event/activity take place, specifically? ( <i>Please do not add any personally-identifiable information, such as addresses or specific names.</i> )	<b>text box</b>
12	event location: <i>regularity</i> <i>(if selected a familiar location for Q10)</i>	How often have you been to this/these place(s) in the past?	<b>multiple choice</b>
			<ul style="list-style-type: none"> <li>• occasionally</li> <li>• frequently</li> </ul>
13	event mood	Please rate how this event/activity made you feel: <ul style="list-style-type: none"> <li>• Positive</li> <li>• Calm</li> <li>• Negative</li> <li>• Afraid</li> <li>• Excited</li> <li>• Happy</li> <li>• Sad</li> </ul>	<b>1-5 scale</b> (for each emotion) 1 = not at all 5 = a great deal
14	event regularity*	Please choose a field to describe the event/activity. This activity is...	<b>multiple choice</b>
			<ul style="list-style-type: none"> <li>• routine (something I do almost every day)</li> <li>• periodic (something I do occasionally)</li> <li>• new (something I've never done before)</li> </ul>
15	event influence from COVID-19	Could you imagine yourself doing this activity: <ul style="list-style-type: none"> <li>• before the COVID-19 pandemic started</li> <li>• at the height of the COVID-19 pandemic</li> <li>• within the next 3 months</li> <li>• a year from now</li> </ul>	<b>1-5 scale</b> (for each time period) 1 = not at all 5 = definitely
16	event memorability	How likely do you think you are to remember this event in the future?	<b>1-5 scale</b> 1 = extremely unlikely 5 = extremely likely
17	event meaningfulness	How meaningful was this event/activity to your life?	<b>1-5 scale</b> 1 = not at all 5 = a great deal
18	event impact	Do you feel that this activity/event will have important consequences for your future?	<b>1-5 scale</b> 1 = not at all 5 = a great deal
19	pre-event experience	What were you doing right before this event/activity happened?	<b>text box</b>
20	post-event experience	What were you doing right after this event/activity happened?	<b>text box</b>

**Appendix C, Table 4.** Memory test questions (including both event recall and temporal memory).

<b>Q#</b>	<b>test item</b>	<b>test prompt</b>	<b>response format</b>
<b>EVENT RECALL</b>			
1	memory vividness	How vivid is your memory of the event titled [event title]?	<b>1-5 scale</b> 1 = not vivid at all 5 = extremely vivid
2	free recall	Recall the event [event title] in as much detail as possible. Describe any and all aspects of the experience that you can remember.	<b>text box</b>
3	event location	Where did this event take place?	<b>text box</b>
4	event location confidence	How confident are you in your answer to the previous question?	<b>1-5 scale</b> 1 = not at all 5 = very much so
5	event date	On which date did this event occur? <i>(enter response as MM/DD/YYYY)</i>	<b>text box</b>
6	event date confidence	How confident are you in your answer to the previous question?	<b>1-5 scale</b> 1 = not at all 5 = very much so
7	event time of day	When during the day did this event occur?	<b>multiple choice</b> <ul style="list-style-type: none"> <li>• morning</li> <li>• afternoon</li> <li>• evening</li> <li>• late night/early morning</li> </ul>
8	event time of day confidence	How confident are you in your answer to the previous question?	<b>1-5 scale</b> 1 = not at all 5 = very much so
9	event social companion(s)	Who were you with when this event happened? <i>(separate individuals by commas)</i>	<b>text box</b>
10	event social companion(s) confidence	How confident are you in your answer to the previous question?	<b>1-5 scale</b> 1 = not at all 5 = very much so
11	event duration	How long did this event last? <i>(enter response as HH:MM)</i>	<b>text box</b>
12	event duration confidence	How confident are you in your answer to the previous question?	<b>1-5 scale</b> 1 = not at all 5 = very much so
13	pre-event experience	What did you do right before this event?	<b>text box</b>
14	post-event experience	What did you do right after this event?	<b>text box</b>
<b>TEMPORAL MEMORY TEST</b>			
0	event pair prompt	Consider the events titled [event 1 title] and [event 2 title].	
1	temporal order	Which event was experienced earlier in time?	<b>multiple choice</b> <ul style="list-style-type: none"> <li>• [event 1 title]</li> </ul>

---

		• [event 2 title]
2	temporal order confidence	How confident are you in your ranking? <b>1-5 scale</b> 1 = not at all 5 = very much so
3	subjective temporal distance	How far apart in time do these events feel? <b>1-5 scale</b> 1 = very close 5 = very far
4	objective temporal distance	How much time (in number of days) actually passed between these events? <b>numeric input</b>
5	event pair similarity	How similar are these events to each other? <b>1-5 scale</b> 1 = not at all 5 = a great deal
6	memory vividness (event 1)	How vivid is your memory for the event titled [event 1 title]? <b>1-5 scale</b> 1 = not vivid at all 5 = extremely vivid
7	memory vividness (event 2)	How vivid is your memory for the event titled [event 2 title]? <b>1-5 scale</b> 1 = not vivid at all 5 = extremely vivid

---

**Appendix C, Table 5.** Results from Bayesian regression models assessing how memory for autobiographical events is shaped by novelty. Directional posterior probability (*pd*) refers to the proportion of each model term's posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd* > 0.975). Models include random slopes for all main effects, as well as random intercepts for each participant.

event memory ~ all novelty metrics (Bayesian)					
	Fixed Effect	Estimate	Std. Error	95% CrI	pd
<b><i>memory vividness scores</i></b>					
	(Intercept)	3.516	0.138	[3.249, 3.788]	1.000
	event regularity ( <i>periodic vs. new</i> )	-0.691	0.096	[-0.882, -0.503]	1.000
	event regularity ( <i>routine vs. new</i> )	-1.496	0.122	[-1.736, -1.256]	1.000
	novel locations visited	0.280	0.077	[0.127, 0.434]	1.000
	novel social interactions	0.039	0.074	[-0.109, 0.187]	0.699
	day notability	0.265	0.082	[0.099, 0.427]	0.999
	day typicality	-0.055	0.037	[-0.130, 0.017]	0.935
<b><i>number of episodic details recalled</i></b>					
	(Intercept)	6.332	0.479	[5.381, 7.275]	1.000
	event regularity ( <i>periodic vs. new</i> )	-0.465	0.266	[-0.989, 0.070]	0.956
	event regularity ( <i>routine vs. new</i> )	-1.209	0.381	[-1.954, -0.437]	0.998
	novel locations visited	0.487	0.323	[-0.150, 1.135]	0.934
	novel social interactions	0.493	0.246	[0.008, 0.988]	0.977
	day notability	0.128	0.253	[-0.367, 0.630]	0.693
	day typicality	-0.089	0.114	[-0.316, 0.13]	0.783

**Appendix C, Table 6.** Results from a Bayesian regression model in which memory vividness scores (for “routine” and “periodic” events only) were predicted by whether or not the event occurred on the same day as a “new” event. Directional posterior probability (*pd*) refers to the proportion of each model term’s posterior distribution that is in the direction of the fixed effect estimate (positive or negative). Bolded *pd* values denote terms for which the 95% credibility interval (*CrI*) excludes 0 (and, equivalently, where *pd* > 0.975). Model includes random slopes for all main effects, as well as random intercepts for each participant.

**memory vividness ~ novel event on same day (Bayesian)**

<i>Fixed Effect</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>95% CrI</i>	<i>pd</i>
(Intercept)	2.639	0.091	[2.460, 2.820]	1.000
novel event on same day	0.266	0.085	[0.059, 0.392]	0.996