Hippocampal and prefrontal regulation of flexible decision updating during memory-guided learning and decision-making

Jesse Thomas Miles

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Sheri Mizumori, Chair

David Gire, Chair

Elizabeth Buffalo

Program Authorized to Offer Degree:

Neuroscience

©Copyright 2023
Jesse Thomas Miles

University of Washington

Abstract

Hippocampal and prefrontal regulation of flexible decision updating during memory-guided learning and decision-making

Jesse Thomas Miles

Co-chairs of the Supervisory Committee:

Sheri Mizumori & David Gire

Department of Psychology

Decision-making requires us to integrate information that is spread out across timescales ranging from milliseconds to years. The information we use to guide decisions depends on current goals and the context in which the decisions are made, all of which can change within moments. As such, we can expect variations in decision-making behavior to depend on requirements of the tasks we engage in, what we've learned about our current situation, and how flexibly we behave. Studies of decision-making behavior in rodents have documented a process in which subjects visibly vacillate between options before committing to a final decision. The behavior, widely known as vicarious trial and error (VTE), has garnered attention as an outward manifestation of deliberation/indecision, and has been shown to vary with task proficiency and changes in task demands. Interestingly, similar phenomena have been reported in both non-human primates and humans, suggesting VTE and related behaviors are widespread.

Despite almost a century of investigation, what exactly defines a VTE can be unclear, and the field has not yet converged on a method for their identification. Plus, we're still learning how the litany of brain structures underpinning decision-making interact to enable the decision-making process – especially as it varies from trial-to-trial during learning, and as behaviors like VTE occur. Further, we tend to take for granted that decision-making, learning, memory, and behavioral flexibility are associated with one another, but we often can't say whether that association is fundamental or because we've defined them such that they depend on one another.

This thesis addresses these issues by proposing and comparing ways to identify VTE, providing new insights into how neural circuitry that supports learning and memory also enables flexible decision-making and VTE, and re-operationalizing how we assess and compare behavioral flexibility and learning to help define behavioral contexts.

Acknowledgements

Academic Acknowledgements

Even in a field as young as neuroscience, I'm reluctant to concede that anyone accomplishes anything solely on their own. We work in the context of an ever-growing literature that we borrow ideas from, and are trained by experts who familiarize us with paradigms, teach us procedures, and help us think about the systems we work with. With this in mind, I would like to acknowledge the academic help I received during the completion of my graduate work.

* * *

Contributions to chapter 2:

Conceptualization – Jesse T. Miles (JTM), Kevan S. Kidder (KSK), David H. Gire (DHG), Sheri JY Mizumori (SJYM)

Methodology – JTM, KSK, SJYM

Data collection and annotation – JTM, KSK, Yiru (Stephanie) Zhu (YZ), Ying Rao (YR), Victoria Hones (VH)

Software – JTM, Ziheng (Christina) Wang (ZW)

Formal analysis – JTM

Writing: original draft – JTM

Writing: editing and revisions – JTM, KSK, DHG, SJYM

Supervision – DHG, SJYM

Contributions to chapter 3:

Conceptualization – KSK, JTM, Phillip M. Baker (PMB), DHG, SJYM

Methodology – KSK, JTM, PMB, SJYM

Data collection and curation –KSK, JTM, VH

Software – KSK, JTM

Formal analysis – KSK, JTM

Writing: original draft – KSK

Writing: editing and revisions – KSK, JTM, VH, PMB, DHG, SJYM

Supervision – SJYM

Contributions to chapter 4:

Conceptualization – KSK, JTM, SJYM

Methodology – KSK, JTM, SJYM

Data collection and curation –KSK, Ryan M. Gillis (RMG), Maeve Bottoms (MB)

Software – RMG, JTM

Formal analysis – KSK, RMG, JTM

Writing: original draft – KSK, RMG

Writing: editing and revisions – KSK, RMG, JTM, SJYM

Supervision – SJYM

Contributions to chapter 5:

Conceptualization – JTM, SJYM

Methodology – JTM, SJYM

Data collection and curation –JTM, Ginger L. Mullins (GLM), Arman Khan (AK)

Software – JTM

Formal analysis – JTM

Writing: original draft – JTM

Writing: editing and revisions –JTM, SJYM

Supervision – SJYM

* * *

Personal Acknowledgements

I was also fortunate to have had excellent scientific training prior to my Ph.D. work, during both undergrad and at the Allen Institute for Brain Science. While in undergrad, Dr. Marti Bosma was kind enough to grant me a position in her lab essentially based on enthusiasm alone, and while there I was lucky to have received additional training from Dr. Hirofumi (Hiro) Watari and Dr. Amanda Tose. It was a wonderful lab environment to learn in, and I'm grateful for the enthusiastic and welcoming introduction to science.

Dr. Jack Waters helped me expertly manage my time across the multiple projects that I contributed to and led at the Allen Institute. I would be lucky to recreate the efficiency in communication that Jack and I were able to strike in my own mentor-mentee relationships, and will strive to balance hands-off and active management as skillfully as Jack did with me. Additionally, much of my success and training at the Institute was due to the generous and helpful supervision of Drs. Matthew Valley, Rylan Larsen, Jun Zhang, and Doug Ollerenshaw. I hope to be as kind and patient with trainees as Matt, Rylan, Jun, and Doug were with me.

* * *

During grad school, help has come from all directions. I am so grateful to the undergraduate research assistants I've been fortunate to work with. Stephanie Zhu, Caroline Jones, Christina Wang, Ryan Gillis, Ginger Mullins, Trinity Charles, Arman Khan, and Maeve Bottoms – you've all been so helpful, and I am looking forward to the bright futures you all have in store. You have been wonderful colleagues.

My lab-mates have been unending sources of insight and relief. Dr. Kevan Kidder has been my closest collaborator, who has kept me honest, excited, and engaged. I'm glad he's been the first to share my successes, failures, breakdowns, and breakthroughs. Victoria Hones has been an exceptional collaborator and source of encouragement, solidarity, and positivity. Her perseverance is exemplary. It's been gratifying to see her emerge as an expert in her field, and a delight to work with her. Dr. Brian Jackson has always been willing to lend an ear, even halfway across the world. His friendship in and out of lab has gone a long way in making grad school tenable, and I'm thankful for his ability to keep me grounded. Dr. Mohammad Tariq has been a critical and keen sounding board for harebrained and intriguing ideas alike (as long as I'm willing to let him workshop a joke). Dr. Suzanne Lewis has been inspiringly tenacious, and has kept me on my toes with her incisive questions and thoughtful discussions.

Of course, the majority of my success in graduate school can be attributed to my thesis advisors and graduate mentors. Dr. Phil Baker ensured we had a clear path forward before leaving the Mizumori lab, and I'm thankful for his ability to always leave me feeling a bit better off after our discussions. Dr. David Gire has been an excellent listener and thinker whose mentorship has been equally helpful in his office, over Zoom, or at the Pub. David's guidance has helped me bridge literatures in a way that I hope emphasizes the bigger picture in our work, and I'm grateful for his careful redirections that have kept me on track. Finally, Dr. Sheri Mizumori has been the most kind, professional, and supportive mentor one could ask for. Never afraid to pull back the curtain, I've learned almost as much about the machinations of academia from Sheri as I have about hippocampal physiology, behavioral neuroscience, and systems neuroscience. A rare combination of steadfastness and open-mindedness, Sheri is a model for how to be a diplomatic and inclusive scholar.

*　　*　　*

To my friends and family – If you're reading this, I appreciate you so much, but I just don't have the space or emotional bandwidth to thank you properly here. Hold me to it, though, and know that I could not have done it without you. I will, however, say a (woefully inadequate) thank you to my mom, Tracy Scribner. No one else deserves more credit for getting me to this point than her.

Preface

I’ve worked with a lot of people who study the brain because they, or those they love, have personally experienced what happens when brains start doing things we don’t understand. Considering all of the inspiring and heart-warming reasons many scientists have for entering their field, I’m often embarrassed at my reason for developing an interest in neuroscience. I was just daydreaming during the first day of my biology class, unable to answer the seemingly straightforward question of: “What is a thought?” I landed on something along the lines of: “Chemicals and electricity moving through the brain,” which did not feel satisfying or well-informed.

A couple years later I answered an email from a student who wanted to start a neuroscience publication at the University of Washington. We called it *Grey Matters*, and I found myself much more interested in devoting my time to the publication than to classes. I read papers way beyond my comprehension, tried to sort out what I’d learned into articles we could publish, and hardly concealed my excitement for our small group’s weekly meetings. My first attempt at an article about large-scale neuroplasticity was pretty thoroughly destroyed by a much more knowledgeable student. Though perhaps a bit harsh, they were fair in their criticism that much of what I was writing was not well aligned with the scientific consensus of the time. So instead, I wrote about the more widely held belief that certain brain areas are specialized for processing certain types of information. Although I still think this was the right thing to have done, looking back, I wish I’d paid more attention to my initial insights which, as I have learned more about the multi-modal and distributed nature of neural information processing, weren’t so far afield from emerging views. At the time, though, I didn’t have the scientific literacy or communication skills to convey this. Just a lot of excitement and curiosity.

Equipped with a better appreciation for the complexity of the terrain, and how much I needed to learn, I decided to follow the lead of my peers and try looking for first-hand research in a neuroscience lab. There are many things to say about these early years in my training, but the most telling is that I’m now writing a thesis on “my contributions” to the field of neuroscience. I fell in love with the bench work, the analysis, the thrill of discovery, and the complexities of the problems. And I’ve never forgotten that first question, which, as far as I know, still doesn’t have a particularly satisfying answer. (If pressed, I’m sure some people could draw you a crude map of the neural regions most heavily implicated in the process of “thinking”, but I suspect there would be a good deal of disagreement, or criticism that the question was ill-posed in the first place, which I’d say is a fair point, but one that highlights the chasm between what we can say as experts and what other folks might expect us to know.)

So, after over ten years of unsuccessfully answering my initial question, I can see why people in my life can be wary of why I’m still doing this, or why I started in the first place. That’s part of why I want to include these personal anecdotes in the preface to what is supposed to be a scholarly work. (That, and because I’ve been told that no one reads your thesis, so what’s the harm?) I regularly wish I had explained myself a little better when asked what I’m doing, why I’m doing it, and what the point is. I’m lucky that much of the time, people are willing to buy into my enthusiasm about the (sometimes tenuous) links between the basic science I do and its applied aspects that might be fruitful in the near future. But when my enthusiasm wanes, or when people question why I’m not working to directly understand the diseases and disorders of the nervous system that they are painfully familiar with, my reasons feel inadequate. I've not done a good job at explaining how I got here, I think, because I struggled with the notion that my own personal interest and aptitude are sufficient motivation. But people become pilots, mechanics, chefs, carpenters, *etc*. for the same reasons – they find something that holds their attention and they become good enough at it to make a living. I'd like to convince myself these simple reasons are good enough, and hope those who haven't always liked my answers will give them consideration as well.

*   *   *

In what follows, I will not be answering my initial question of "What is a thought?". Instead, I'll be looking at how decisions are made when recent experience dictates how to act. I'll tackle this using rats, which allows me to examine a well-known but relatively understudied decision-making behavior that looks like a real-time change of mind, while simultaneously recording, or, in some cases manipulating, neural activity. What I've (or, more accurately, *we*'ve) found in this work doesn't necessarily speak toward what a thought *is*, or where it resides, or how it travels through the brain. But I'd like to think it gives us some clues as to how uncertainty and learning shape our behavior, and helps us understand which specific brain regions allow for the proper expression of indecision or deliberation as we learn to let memories guide our decision-making. Ask me about what "thoughts" are in another decade or so

Table of Contents

Chapter 1

Introduction

1.1 The (intuitive) link between learning and decision-making

Not all decisions are the same. It's typically easier to decide what to have for dinner than how to advance your career, or to choose your next move in a game you've played before compared to one you are still learning. Similarly, we can appreciate the difference between choosing to stop when you encounter traffic on a normally uncongested street compared to choosing an alternative route when you see traffic ahead. Despite these intuitions, much of the research devoted to understanding what's going on in the brain as we make decisions or develop strategies is not designed to understand the nuance involved in the choices we make.

From a neuroscience standpoint, much of what we know about the neural underpinnings of learning and decision-making comes from studies in rodents because we can carefully design experiments where we manipulate, and/or record activity from, the brain in ways we usually cannot in people. One difficulty this introduces, of course, is the inability to ask what drove a subject's choice, leaving us with the non-trivial problem of inferring a rat's rationale and learning based on behavioral and neural patterns. To codify these behavioral patterns, studies of learning and decision-making often employ tasks that have multiple trials, each with a correct answer, based on rules that are simple enough to learn through trial and error. In the early half of the 20th century, this included tasks that required subjects to: discriminate between corridors with different visual cues (Muenzinger, 1938; Yerkes, 1907) or between tones of different frequencies (Muenzinger, 1938), determine whether reward would be delivered at a particular location or in response to a particular movement direction (Tolman et al., 1946), or navigate through complex mazes to find rewards (Hubbert & Lashley, 1917; Tolman, 1948). The important point is that after we linked learning to accurate decision-making, we took that link for granted.

Today's behavioral neuroscience paradigms look largely like they did in the early 20th century. We still tend to rely on repeated, accurate decisions as signs that decision-makers have (ostensibly) learned the rules of a task. We've certainly made strides in understanding learning, memory, and decision-making by relying as heavily as we have on this single behavioral measurement, but it means that we've almost inextricably tied learning to accurate decision-making. The overarching theme of this thesis is a discussion and demonstration of how to go beyond simple choice outcome readouts to help understand how the brain supports flexible memory usage to guide decision-making. First, though, we'll need to discuss what it might look like for a rat to "think".

1.2 Variations in decision-making behavior(s)

Evelyn Gentry and Karl Muenzinger (and others before them e.g. Yerkes, 1907, p. 130, 224) had noticed that rats did not always immediately indicate their decision when presented with multiple options. Instead, they would sometimes reach a critical juncture and appear to consider the possibilities before making their final choice. Gentry's 1930 master's thesis explicitly studied this behavior, which she and Muenzinger called “vicarious trial and error”, proposing it as a proxy for actual trial-and error behavior. The idea was that instead of committing to a choice and experiencing its consequences, rats would simulate and compare the expected outcome of alternatives before their ultimate decision, hence the “vicarious” in vicarious trial and error (hereafter referred to as VTE).

Gentry and Muenzinger’s early observations on VTE can be summarized by saying that VTE appeared to be more likely during learning, and that more efficient learners tended toward higher rates of VTE (Muenzinger, 1938; Muenzinger & Gentry, 1931). Though he didn’t use the term VTE, a more provocative description of the behavior came from Tolman (1926) who likened the act of hesitating and alternately considering options at choice points to conscious awareness. The primary similarity in Gentry, Muenzinger, and Tolman’s early thoughts on VTE are captured by Tolman’s assertion that conscious awareness required “a *representation* of results so that the latter can themselves become determiners for or against the act which leads to them,” (Tolman, 1926, p. 366). Put another way, if decision-makers can consciously represent the outcome of a behavior before carrying it out, they can then decide what to do based on whether the represented outcome was favorable or not. Crucially, outcomes can only be represented/simulated/imagined (choose your favorite term) if one has learned some cause-and-effect relationship between actions and outcomes. After forming these associations, you can utilize prior memories, or what you’ve learned about the rules governing your circumstance, to guide upcoming decisions.

This insight was crucial in Tolman’s hypothesis that “in the course of learning, something like a field map of the environment gets established,” (Tolman, 1948). This map was quite loosely described by Tolman as a feature of the nervous system that enables representation of related environmental features. In proposing the so-called “cognitive map”, Tolman was expressing his disagreement with the widely held views that actions were strictly dictated by sensory stimuli, and that the formation of associations between sensation and action could be explained entirely by repeated stimulus-response pairing. Tolman pointed out that because VTE tended to occur around crucial parts of the learning and decision-making processes, and because VTE rates could change even if sensory stimuli did not, it was unlikely that they reflected a passive, stimulus comparison process that led to a determined response. Instead, he saw them as an active process where events were compared and evaluated before a response. In the parlance of a (cognitive) map, “it is this tentative map, indicating routes and paths and environmental relationships, which finally determines what responses, if any, the animal will finally release,” (Tolman, 1948, p. 192).

Accordingly, the most popular modern treatment of VTE proposes that VTE might serve as a behavioral marker of deliberation (Redish, 2016), a notion that has received considerable support (Bett et al., 2012; Blumenthal et al., 2011; Hasz & Redish, 2020; Papale et al., 2012, 2016; Schmidt et al., 2013, 2019; Stout et al., 2021). This idea is a slight rephrasing of Tolman’s formulation but is similar in that deliberation is meant to highlight a serial search through possibilities, which requires an understanding of which outcomes are likely or possible, and how they relate to available choices. In this sense it is also compatible with the classic idea that VTE should accompany learning because some connection between possibilities and choices must be learned for any evaluation and comparison to occur (even if the mapping between choices and outcomes proves incorrect). It’s worth emphasizing, though, that VTE is not the only possible indicator of deliberation, and that deliberation may occur without VTE. Thus, the presence or absence of one is not necessarily indicative of the presence or absence of the other. The broader point, however, is that VTE indicates an interesting variation in decision-making behavior that we may also be able to use in addition to choice outcome when assessing learning and memory.

The case made so far suggests that VTEs could be valuable in helping us understand memory-guided decision-making, highlighting that the link between the decisions and memory might be reliant on a "cognitive map" relating options to learned possibilities. Thinking back to an example given earlier, traffic only affects my decision-making if I can remember an alternate route. As such, we might expect the same brain structures involved in learning and memory to be involved in VTE generation and regulation. The next sections will describe one of those structures, the hippocampus, and explain how patterns of hippocampal neural activity have contributed to the hypothesis that it could act as the brain's basis for Tolman's cognitive map.

1.3 Hippocampal physiology – memory, maps, and sequences

It wasn't until the mid-1950s that a series of clinical procedures and psychological assessments in patients suffering from intractable epilepsy and severe behavioral control issues demonstrated a critical role of the hippocampus in learning and memory. Dr. Brenda Milner, working with neurosurgeons Wilder Penfield and William Scoville, demonstrated that patients who had undergone experimental removal of the hippocampi had extreme difficulty remembering recent events (Milner, 1954; Milner & Penfield, 1955; Scoville & Milner, 1957). Around the same time, procedures and technologies developed that enabled recording from the behaving brain. In a seminal set of experiments, Green & Arduini, (1954) recorded from a variety of animals under all sorts of experimental conditions to describe how patterns of hippocampal and cortical local field potential (LFP) patterns related to different sensory stimuli, stimulation or lesion of certain brain regions, and each other. In a statement that seems to have anticipated the coming 70 years of research, they mention that the hippocampus "may well share, with the frontal cortex, many important functions of a delicacy which is not readily susceptible to analysis", (Green & Arduini, 1954, p. 551). At least part of this conclusion appears driven by the fact that hippocampal activity they saw persisted without much interruption after removing the neocortex, and that gross motoric and affective behaviors appeared relatively unchanged after removal of the hippocampus in other experiments.

Since these early investigations, we have indeed learned that the correlates of hippocampal activity are often "delicate" and difficult to analyze. Early studies looking at activity in individual hippocampal neurons discovered a remarkable connection between activity of certain cells and particular spatial locations (O'Keefe & Dostrovsky, 1971; O'Keefe & Nadel, 1978; Ranck, 1973). All of these researchers noted that as rats foraged for treats, certain cells (now known as "place cells") would become increasingly active as rats passed through a particular location (their "place field"). Building on Tolman's purely hypothetical model of a cognitive map, and behavioral experiments in which the hippocampus was lesioned, O'Keefe and Nadel formalized the observation of place-reactive cells into the proposal that the hippocampus can act as the anatomical and physiological substrate for an allocentric, cognitive map – one that is certainly spatial, given the spatially correlated firing patterns, but could also be semantic (see p. 493 of O'Keefe & Nadel, 1979).

In addition to spatial correlates, which were admittedly found in the minority of hippocampal cells, these reports and others found many types of behavioral and sensory evoked activity in the hippocampus (see especially Ranck, 1973 and Vinogradova, 1970). In fact, activity could be referenced to so many things that Ranck's summary of his 1973 findings left the door pretty wide open in terms of describing hippocampal function based on cellular activity correlates:

> "Hippocampal transformations would seem to help solve such problems as how to sequence various automatic behaviors appropriately, how to sequence automatic and nonautomatic behaviors appropriately, how to test the appropriateness of an automatic behavior or sequence and stop or change it if needs be, how to shift from one behavior to another, how to combine automatic and non-automatic behaviors into new patterns, how to use behaviors which are being learned along with older behaviors, or in general how to use these automatic and nonautomatic behaviors in a flexible way and to avoid being too rigid," (Ranck, 1973, p. 523-524).

This characterization makes it clear that the observed generality of hippocampal activity did not make it amenable to simple descriptions, particularly if one wanted to account for activity not directly explained by the spatial framework of the cognitive map hypothesis.

A different characterization by O. S. Vinogradova, who described how repeated sensory stimuli changed hippocampal activity, proposed that the hippocampus was crucial in comparing new and old signals. Based on a simple behavior known as the orienting reflex, where subjects behaviorally indicate that they perceive some stimulus (e.g. by turning toward a sound source), Vinogradova showed that, just as the orienting reflex diminished after multiple presentations of the same, neutral (i.e. unrewarded, unpunished, and otherwise non-salient) stimulus, activity in hippocampal neurons would often decline in response to repeated presentations of the same stimulus (Vinogradova, 1970). In doing so, she also showed that hippocampal neurons could develop anticipatory firing with respect to the timing of stimuli, as long as the timing was scheduled predictably. As expected, this anticipatory firing would fade after many presentations. Equally important, presentations of a new stimulus would often lead to increased activity in neurons that had previously stopped or decreased their responses to the old stimulus. These observations were summarized by concluding that the main function of the hippocampus was to detect novelty by determining whether something matched prior experience, a process which could be considered a basic form of memory. Interestingly, Vinogradova did not consider the hippocampus as the source of those stored experiences (the locus of memory), only as the comparator (see p. 115 of Vinogradova, 1970).

Despite the wide variety of observed hippocampal correlates, O'Keefe and Nadel's theory of the hippocampus as a cognitive map seems to have received (and perhaps still receives) the most attention, likely because of the specificity of the framework and detailed synthesis of information to support the theory. Still, although it was not likely their goal to describe the map as purely spatial in nature, critics of the theory latched onto the spatial framing as a way to poke holes in it. A succinct, yet pointed, critique came from Eichenbaum and Cohen, (1988), who described a series of experiments finding activity in hippocampal units during explicitly non-spatial tasks. And while they do not try to argue against the hippocampus as having spatial correlates, they instead propose that:

> "A more appropriate description would be that hippocampal complex-spike neurons represent various relationships among multiple stimuli and contingencies or responses, including configurational properties of items simultaneously present in the environment and significant relationships between temporally separated items," (Eichenbaum & Cohen, 1988 p. 246).

This formulation posits that hippocampal activity could be thought of as representing the relationships between physical, external objects, just as in the spatial interpretation, but also that it could represent abstract relationships between objects in memory. It also rings true to many of Ranck's proposals, particularly those that have to do with the appropriate sequencing of related behaviors, and Vinogradova's observations that many different stimuli could alter hippocampal activity – especially when presented at unpredictable timing intervals. In all cases the crucial element uniting descriptions of hippocampal activity relied on how stimuli or behaviors related to one another.

Regardless of whether space was the only, or even primary, type of relationship that the hippocampus could be shown to represent, the spatial framing did enable several impactful discoveries. Several reports demonstrated that cells which had been identified as place cells during spatial reasoning or exploration tasks could also reactivate during sleep (Pavlides & Winson, 1989; Wilson & McNaughton, 1994), and that these reactivations during sleep were preferentially ordered in sequences that mimicked their temporal order during exploration (Lee & Wilson, 2002; Skaggs & McNaughton, 1996). In support of the notion that hippocampal activity may only be incidentally associated with space, another group published similar findings of highly specific, sequenced activity that had no explicitly spatial correlates (Nádasdy et al., 1999).

All of these findings were critical for their recognition that relational information coding, be it spatial or otherwise, ought to show up as orderly, sequenced activity. What's more, they helped bridge a gap between hippocampal physiology, clinical observations (e.g. Milner, 1972; Scoville & Milner, 1957; Zola-Morgan et al., 1986), and behavioral results (Aggleton et al., 1986; Fortin et al., 2002; Morris et al., 1982; O'Keefe & Nadel, 1978; Olton et al., 1978) implicating the hippocampus in learning and memory. These results also demonstrated that, during sleep, the sequences tended to be strongly enriched during high frequency hippocampal LFP events called sharp-wave ripples, which had been hypothesized as a mechanism for information transfer from the hippocampus to the neocortex (Chrobak & Buzsáki, 1994, 1996; Wilson & McNaughton, 1994).

Taking stock of these results from the latter half of the 20th century, the emergent picture was that of the hippocampus as a cognitive map, but a map in the very general sense that it related things to one another. Its activity could represent diverse elements of the outside world, organized according to their relationships, and with respect to orderly behavioral and anticipatory sensory sequences. These sequences required experience, and, presumably, could transfer to neocortex if consolidated into longer term memory. Less clear, though, was the relationship between decision-making and the working memory that guided those decisions.

1.4 Hippocampal physiology & memory-guided decision-making

Vinogradova's reports in the 1970s demonstrated that hippocampal neurons could anticipate, and habituate to, regularly timed sensory stimulation. In other words, it was clear that hippocampal cells could have some sensitivity to prior and upcoming (sensory) events. In a similar vein, experiments that recorded from the same putative place cell during different tasks and in different environments showed that cells could change their firing characteristics with respect to prior or upcoming movements, or even in response to abstract task demands in the same environment (Markus et al., 1995). As additional reports showed that some hippocampal place cell activity was sensitive to recent or upcoming events (Ferbinteanu & Shapiro, 2003; Frank et al., 2000; Griffin et al., 2007; Wood et al., 2000), it became clear that at least some place cells were not necessarily active *any* time a rat occupied or entered a certain location. Sometimes, they were only active if they had also come from a particular location, were about to go to a particular location, or some ambiguous combination of the two.

These results contradicted what would be expected from purely spatial cells that ought to fire any time a particular position is occupied, regardless of what has or is about to happen. Instead, they were interpreted as indicators of memory episodes or demands (Ferbinteanu & Shapiro, 2003; Griffin et al., 2007; Wood et al., 2000), ways to situate behaviors within neural reference frames (Markus et al., 1995), or neural substrates that could support planning by processing or representing information from the recent past, present, and possible future (Frank et al., 2000). This latter interpretation was further supported by the identification of place cell sequences reflecting the prior, current, and future position of rats (Dragoi & Buzsáki, 2006; Foster & Wilson, 2007).

Since VTE had been proposed as a candidate behavior linked to planning with a cognitive map – or, if nothing else, a clear variant in decision-making behavior – a natural question was whether there was a relationship between VTE and hippocampal activity or manipulation. Hu and Amsel, (1995) showed that, compared to control rats with hippocampi intact, rats with hippocampal lesions performed fewer VTEs and did not learn a visual discrimination task as quickly (if they learned it at all). Remarkably, Johnson and Redish (2007) connected the observation of non-local, future-oriented hippocampal representations to VTE likelihoods and error corrections. They showed that trajectories where rats changed their orientation of motion (indicative of head reorientations) were correlated with increases in non-local hippocampal representations. They also showed that some error correction trials were associated with hippocampal place cell "sweeps" of sequences that first traced the alternate possibility *before* the behavioral correction. Another notable observation was that, compared to the typical place cell sequences that had been shown to be organized within high-frequency sharp-wave ripples, non-local sequences reported here occurred during times when the LFP contained very different types of activity in much lower frequency bands. Intriguingly, one of these lower frequency LFP bands, the roughly 6 to 12 Hz theta band, had already been reported to bracket place cell sequences (Dragoi & Buzsáki, 2006; Foster & Wilson, 2007). That said, many of the sequence alternations seen by Johnson and Redish unfolded over longer timescales. Still, these results were the most direct evidence so far of the relationship between hippocampal activity and memory-guided decision-making.

1.5 Interactions with the prefrontal cortex during decisions

As the field worked toward reconciling the seemingly disparate accounts and theories of hippocampal function, researchers also began exploring an interplay between the hippocampus and a section of cortex known as the prefrontal cortex (PFC, or, specifically, the medial (m)PFC in rodents). Reports from PFC of monkeys showed that activity could span the delays of working memory tasks (Fuster & Alexander, 1971; Kubota & Niki, 1971), and reports in rats showed that the mPFC also contained spatial correlates, though these spatial correlates seemed more related to task phases than space *per se* (Jung et al., 1998). Furthermore, both monkeys (Goldman-Rakic, 1991) and rats (Seamans et al., 1995) showed difficulty in spatial and working memory tasks when the PFC was lesioned. Since many of these observations mirrored what had been seen when recording from or lesioning the hippocampus, it seemed clear that information was being shared between regions.

One clever strategy to specifically show that *interactions* between the two areas mediated behavioral performance was to reversibly silence each region in one hemisphere, but on opposite sides of the brain. In this way, processing by each region functioned normally on one side of the brain. However, since these regions were only anatomically connected in the same hemisphere, information could not be shared between areas. Using this “contralateral disconnection” approach, Floresco *et al.* (1997) were able to demonstrate that rats made more errors in a task that required them to visit locations they had not previously visited. Crucially, these errors were only more likely when the hippocampus and mPFC were disconnected. Interestingly, the errors only occurred if there was an intervening delay between the period when they sampled initially rewarded locations and the period when they were supposed to forage for newly rewarded locations. Additionally, the types of errors were roughly equally distributed between entries into locations visited prior to the delay, and areas already visited while searching after the delay. The authors interpreted these results as a failure to prospectively organize behavior when the hippocampal and prefrontal regions were disconnected. Similar results were reported during a different spatial task that required maintaining memory of a previous choice over a short delay interval, but these authors concluded that the disruption was due to impairments to spatial working memory instead of prospective planning (Wang & Cai, 2006).

Simultaneous recordings from the mPFC and hippocampus (HPC) during a variety of tasks requiring working memory and memory-guided decision-making have demonstrated that these two structures dynamically interact with one another throughout the decision-making process. Many of the studies agree that coherent local field potential (LFP) oscillations, as well as cross-regional alignment of spiking in one area to the LFP of another, show relationships to behavior and task structure. Specifically, these cross-regional interactions tend to occur within the roughly 6 to 12 Hz theta band, the frequency range shown to contain place cell sequences (Dragoi & Buzsáki, 2006; Foster & Wilson, 2007). There also tends to be consensus that HPC-mPFC theta LFP coherence increases as rodents approach locations where they need to physically choose between different paths to reward (Benchenane et al., 2010; Hallock et al., 2016; Jones & Wilson, 2005; Sigurdsson et al., 2010; Tavares & Tort, 2022) away from punishment (Dickson et al., 2022), or out of an anxiety inducing area (Adhikari et al., 2010).

Additionally, when studies have explicitly compared HPC-mPFC theta interactions during correct navigation toward a reward vs incorrect navigation, researchers find increased theta coherence on correct trials compared to error trials, regardless of whether they report field potential coherence or spike phase coherence between units of the mPFC and theta in the HPC (Hyman et al., 2010; Jones & Wilson, 2005). This is often interpreted as evidence of ensembles in both regions synchronously integrating information from local processing in order to guide behavior (Benchenane et al., 2010; Hyman et al., 2010; Jones & Wilson, 2005). A related finding is that coherence during the act of decision-making is higher after a rule has been learned, which is typically defined by increased (or perfect) decision-making accuracy (Benchenane et al., 2010; Dickson et al., 2022; Hallock et al., 2016). A notable exception, however, comes from Guise and Shapiro, who observe increased HPC-mPFC theta coherence just *before* animals embark to make a decision, as well as when they enter and move through a goal location (Guise & Shapiro, 2017).

In addition to interactions at the level of the LFP, which is thought of as a network-level phenomenon, studies on single-cell correlates have found coordinated spatial representations in the HPC and mPFC (Zielinski et al., 2019). Similar correlations in spatial representations between HPC and mPFC have been confirmed by Hasz and Redish, who also report an increased likelihood of representing non-local and putatively prospective spatial information in both areas during VTE (Hasz & Redish, 2020). Interestingly, both of these reports divided their time windows for decoding ensemble activity into cycles of the hippocampal theta oscillation. A recent report showed that the mPFC also contains theta-based sequencing of spatial activity, with these sequences preferentially encoding locations that sweep from behind to ahead of the current position (Tang et al., 2021). These theta sequences tended to be more precise in the hippocampus, while mPFC was more likely to represent the upcoming decision.

Evidence up to this point suggests that information is shared between the hippocampus and mPFC during memory-guided decision-making. Various forms of activity at the single cell, small ensemble, and network level have all been implicated in supporting the process. In particular, many studies seem to implicate theta-related activity between the mPFC and hippocampus as playing a crucial role in memory-guided decision-making. One strong test of theta's importance recently came from an experiment that required rats to wait until their mPFC-HPC theta coherence reached above a certain threshold before they were able to make decisions. The authors predicted that, if given the opportunity to choose while the mPFC-HPC network was in a state known to be related to successful decision-making, the performance would be better than usual, which was indeed what they saw (Stout et al., 2023). Furthermore, even though these trials were triggered based on theta coherence at a location distant from the choice point in both space and time, post-hoc analysis showed that theta coherence peaked again just before reaching the choice point on these same trials. And while hippocampal theta did not change significantly during either trial type, mPFC theta power was substantially higher on the trials triggered by high coherence. Taken together, studies that record from both the hippocampus and mPFC suggest that their interactions are closely linked to memory-guided decision-making, and that coordinated, rhythmic activity may be a crucial part of these interactions.

1.6 Relationships between the hippocampus, mPFC, and VTE

Several reports have shown relationships between the hippocampus, mPFC, and VTE since the early findings that hippocampal sequences were more likely to represent distal locations during VTE. Animals with hippocampal lesions are less likely to VTE early in a session, while they are learning reward contingencies for spatial tasks (Bett et al., 2012). High frequency hippocampal ripple events at reward points were shown to be less frequent before trials with VTE, and VTEs were less frequent just before trials with ripples at the reward (Papale et al., 2016). Deformations of the hippocampal oscillation and relative phase relationships between its frequency bands were redistributed during VTEs (Amemiya & Redish, 2018). Disrupting mPFC activity over long time periods decreased the likelihood that VTE occurred, and concurrently altered the typical organization of hippocampal oscillations, as well as the timing and content of theta sequences during VTE (Schmidt et al., 2019). Bilaterally silencing a cortical pathway that links the hippocampus and mPFC increases the likelihood of VTE during periods when it is typically decreased (Kreher et al., 2019). Similarly, silencing a thalamic pathway that connects the hippocampus and mPFC also increases VTE and error rates due to inflexible behavior, and decreases theta coherence between structures during VTEs on incorrect trials (Stout et al., 2021). Silencing the dorsal hippocampus led to delays in both behavioral shifts to a new reward contingency and the onset of VTE behaviors (Meyer-Mueller et al., 2020).

What these results suggest is that when neural systems are not interrupted, VTE behavior typically aligns with, and is perhaps an indicator of, flexible behavior. However, disrupting either the hippocampus, mPFC, or regions that connect the two tends to break this relationship – VTEs will no longer occur with respect to changes in behavior driven by changes in task demand. In delay discounting tasks, for example, subjects usually adjust the delay times between different reward sizes to a certain threshold, and then exploit that threshold. VTE tends to increase during that adjustment period, and then go away as they enter an exploitation period (Papale et al., 2012, 2016). Disrupting the perirhinal link between the hippocampus and mPFC, however, disrupts this typical pattern – rats continue to adjust and VTE well into the second half of the task, which is when they would have usually been able to exploit their preferred delay (Kreher et al., 2019). Disrupting the mPFC decreased VTE for typically difficult decisions (Schmidt, 2019). When dorsal hippocampi were silenced, the likelihood of VTE decreased during the period of flexibility, but increased after a new reward contingency had been learned, which was opposite the relationship observed during typical learning with intact hippocampi (Meyer-Mueller et al., 2020). Silencing the thalamic relay between the dorsal hippocampus led to increases in perseverative errors and corresponding increases in VTEs during those inflexible periods (Stout et al., 2021).

In other words, manipulating the mPFC, HPC, or systems that connect those structures seems to switch VTEs from a behavior associated with deliberation to one associated with indecision, or decrease the likelihood that they will happen at all. If this is indeed the case, then we should be able to demonstrate that VTE can become uncoupled from the learning process during indecision or inflexible decision-making. We should also expect that manipulating these structures can change the relationship between flexible decision-making, choice accuracy, and VTE.

1.7 Summary of the following chapters

The rest of this thesis uses observations from this introduction to further our understanding of how the HPC, mPFC, and behaviors are linked within the context of flexible, memory-guided decision-making. After finding that the most commonly used method of identifying VTE was not reliable in our experiments, Chapter 2 defines a set of simple values that quantify trajectory shapes, and, using common machine learning algorithms, shows that using those values led to more precise and reliable VTE classification. We also go beyond the typical reports of statistical differences between elements of the hippocampal oscillation during VTE by showing that those elements, as well as frequency domain representation of the oscillation, enable above chance predictions of which trials have VTE.

Chapter 3 builds on our knowledge of the mPFC's role in spatial working memory and VTE by showing that brief mPFC disruption during decision-making can decrease choice accuracy, and that brief disruptions across behavioral epochs in general decrease VTE rates. Furthermore, we show the decrease in VTE rates was accompanied by corresponding decreases in choice accuracy. Using the same mPFC disruption paradigm in a task with changing spatial reward contingencies, Chapter 4 shows that the mPFC is causally involved in different working memory process as task demands change, and that unperturbed mPFC function allows for expression of VTE during periods of learning and flexibility. Disrupting the mPFC, however, breaks this relationship without decreasing VTE rates.

Utilizing a task that requires subjects to serially switch between different spatial decision-making strategies, I explicitly test the hypothesis that VTEs are related to learning and flexible behavior. To do so, I formally estimate learning by modeling likelihoods of different behavioral strategies and show that VTE is strongly, but briefly, aligned to task learning. Furthermore, I operationally define behavioral flexibility in relation to changes in strategy likelihoods, and show that increases in flexibility are aligned to both rule learning and changes in VTE rates. Similarly, we show that when we split VTEs by flexibility, more flexible VTEs are more likely to be correct, and correct VTEs are likely to have higher flexibility scores. These observations support the hypothesis that VTEs are typically indicate deliberation (at least in our task), but their increased likelihood of leading to incorrect choices during inflexible periods suggests that they may also be indicative of uncertainty or indecision during these times.

Taken together, this work strengthens our understanding of how variations in decision-making behavior rely on the hippocampus, mPFC, learning, and flexible memory use.

1.7 References

Adhikari, A., Topiwala, M. A., & Gordon, J. A. (2010). Synchronized Activity between the Ventral Hippocampus and the Medial Prefrontal Cortex during Anxiety. *Neuron*, *65*(2), 257–269. https://doi.org/10.1016/j.neuron.2009.12.002

Aggleton, J. P., Hunt, P. R., & Rawlins, J. N. P. (1986). The effects of hippocampal lesions upon spatial and non-spatial tests of working memory. *Behavioural Brain Research*, *19*(2), 133–146. https://doi.org/10.1016/0166-4328(86)90011-2

Amemiya, S., & Redish, A. D. (2018). Hippocampal Theta-Gamma Coupling Reflects State-Dependent Information Processing in Decision Making. *Cell Reports*, *22*(12), 3328–3338. https://doi.org/10.1016/j.celrep.2018.02.091

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., & Wiener, S. I. (2010). Coherent Theta Oscillations and Reorganization of Spike Timing in the Hippocampal-Prefrontal Network upon Learning. *Neuron*, *66*(6), 921–936. https://doi.org/10.1016/j.neuron.2010.05.013

Bett, D., Allison, E., Murdoch, L. H., Kaefer, K., Wood, E. R., & Dudchenko, P. A. (2012). The neural substrates of deliberative decision making: Contrasting effects of hippocampus lesions on performance and vicarious trial-and-error behavior in a spatial memory task and a visual discrimination task. *Frontiers in Behavioral Neuroscience*, *6*. https://doi.org/10.3389/fnbeh.2012.00070

Blumenthal, A., Steiner, A., Seeland, K., & David Redish, A. (2011). Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task. *Neurobiology of Learning and Memory*, *95*(3), 376–384. https://doi.org/10.1016/j.nlm.2011.01.011

Chrobak, J. J., & Buzsáki, G. (1994). Selective activation of deep layer (V-VI) retrohippocampal cortical neurons during hippocampal sharp waves in the behaving rat. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *14*(10), 6160–6170. https://doi.org/10.1523/JNEUROSCI.14-10-06160.1994

Chrobak, J. J., & Buzsáki, G. (1996). High-Frequency Oscillations in the Output Networks of the Hippocampal–Entorhinal Axis of the Freely Behaving Rat. *Journal of Neuroscience*, *16*(9), 3056–3066. https://doi.org/10.1523/JNEUROSCI.16-09-03056.1996

Dickson, C. R., Holmes, G. L., & Barry, J. M. (2022). Dynamic θ Frequency Coordination within and between the Prefrontal Cortex-Hippocampus Circuit during Learning of a Spatial Avoidance Task. *ENeuro*, *9*(2). https://doi.org/10.1523/ENEURO.0414-21.2022

Dragoi, G., & Buzsáki, G. (2006). Temporal Encoding of Place Sequences by Hippocampal Cell Assemblies. *Neuron*, *50*(1), 145–157. https://doi.org/10.1016/j.neuron.2006.02.023

Eichenbaum, H., & Cohen, N. J. (1988). Representation in the hippocampus: What do hippocampal neurons code? *Trends in Neurosciences*, *11*(6), 244–248. https://doi.org/10.1016/0166-2236(88)90100-2

Ferbinteanu, J., & Shapiro, M. L. (2003). Prospective and retrospective memory coding in the hippocampus. *Neuron*, *40*(6), 1227–1239. https://doi.org/10.1016/s0896-6273(03)00752-9

Floresco, S. B., Seamans, J. K., & Phillips, A. G. (1997). Selective Roles for Hippocampal, Prefrontal Cortical, and Ventral Striatal Circuits in Radial-Arm Maze Tasks With or Without a Delay. *Journal of Neuroscience*, *17*(5), 1880–1890. https://doi.org/10.1523/JNEUROSCI.17-05-01880.1997

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), Article 5. https://doi.org/10.1038/nn834

Foster, D. J., & Wilson, M. A. (2007). Hippocampal theta sequences. *Hippocampus*, *17*(11), 1093–1099. https://doi.org/10.1002/hipo.20345

Frank, L. M., Brown, E. N., & Wilson, M. (2000). Trajectory Encoding in the Hippocampus and Entorhinal Cortex. *Neuron*, *27*(1), 169–178. https://doi.org/10.1016/S0896-6273(00)00018-0

Fuster, J. M., & Alexander, G. E. (1971). Neuron Activity Related to Short-Term Memory. *Science*, *173*(3997), 652–654.

Goldman-Rakic, P. S. (1991). Chapter 16 Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. In H. B. M. Uylings, C. G. Van Eden, J. P. C. De Bruin, M. A. Corner, & M. G. P. Feenstra (Eds.), *Progress in Brain Research* (Vol. 85, pp. 325–336). Elsevier. https://doi.org/10.1016/S0079-6123(08)62688-6

Green, J. D., & Arduini, A. A. (1954). HIPPOCAMPAL ELECTRICAL ACTIVITY IN AROUSAL. *Journal of Neurophysiology*, *17*(6), 533–557. https://doi.org/10.1152/jn.1954.17.6.533

Griffin, A. L., Eichenbaum, H., & Hasselmo, M. E. (2007). Spatial Representations of Hippocampal CA1 Neurons Are Modulated by Behavioral Context in a Hippocampus-Dependent Memory Task. *Journal of Neuroscience*, *27*(9), 2416–2423. https://doi.org/10.1523/JNEUROSCI.4083-06.2007

Guise, K. G., & Shapiro, M. L. (2017). Medial Prefrontal Cortex Reduces Memory Interference by Modifying Hippocampal Encoding. *Neuron*, *94*(1), 183-192.e8. https://doi.org/10.1016/j.neuron.2017.03.011

Hallock, H. L., Wang, A., & Griffin, A. L. (2016). Ventral Midline Thalamus Is Critical for Hippocampal–Prefrontal Synchrony and Spatial Working Memory. *Journal of Neuroscience*, *36*(32), 8372–8389. https://doi.org/10.1523/JNEUROSCI.0991-16.2016

Hasz, B. M., & Redish, A. D. (2020). Spatial encoding in dorsomedial prefrontal cortex and hippocampus is related during deliberation. *Hippocampus*, *30*(11), 1194–1208. https://doi.org/10.1002/hipo.23250

Hu, D., & Amsel, A. (1995). A simple test of the vicarious trial-and-error hypothesis of hippocampal function. *Proceedings of the National Academy of Sciences*, *92*(12), 5506–5509. https://doi.org/10.1073/pnas.92.12.5506

Hubbert, H. B., & Lashley, K. S. (1917). Retroactive association and the elimination of errors in the maze. *Journal of Animal Behavior*, *7*(2), 130–138.

Hyman, J., Zilli, E., Paley, A., & Hasselmo, M. (2010). Working memory performance correlates with prefrontal-hippocampal theta interactions but not with prefrontal neuron firing rates. *Frontiers in Integrative Neuroscience*, *4*. https://www.frontiersin.org/articles/10.3389/neuro.07.002.2010

Johnson, A., & Redish, A. D. (2007). Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point. *Journal of Neuroscience*, *27*(45), 12176–12189. https://doi.org/10.1523/JNEUROSCI.3761-07.2007

Jones, M. W., & Wilson, M. A. (2005). Theta Rhythms Coordinate Hippocampal–Prefrontal Interactions in a Spatial Memory Task. *PLOS Biology*, *3*(12), e402. https://doi.org/10.1371/journal.pbio.0030402

Jung, M. W., Qin, Y., McNaughton, B. L., & Barnes, C. A. (1998). Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. *Cerebral Cortex*, *8*(5), 437–450. https://doi.org/10.1093/cercor/8.5.437

Kreher, M. A., Johnson, S. A., Mizell, J.-M., Chetram, D. K., Guenther, D. T., Lovett, S. D., Setlow, B., Bizon, J. L., Burke, S. N., & Maurer, A. P. (2019). The perirhinal cortex supports spatial intertemporal choice stability. *Neurobiology of Learning and Memory*, *162*, 36–46. https://doi.org/10.1016/j.nlm.2019.05.002

Kubota, K., & Niki, H. (1971). Prefrontal cortical unit activity and delayed alternation performance in monkeys. *Journal of Neurophysiology*, *34*(3), 337–347. https://doi.org/10.1152/jn.1971.34.3.337

Lee, A. K., & Wilson, M. A. (2002). Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep. *Neuron*, *36*(6), 1183–1194. https://doi.org/10.1016/S0896-6273(02)01096-6

Markus, E. J., Qin, Y. L., Leonard, B., Skaggs, W. E., McNaughton, B. L., & Barnes, C. A. (1995). Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *Journal of Neuroscience*, *15*(11), 7079–7094. https://doi.org/10.1523/JNEUROSCI.15-11-07079.1995

Meyer-Mueller, C., Jacob, P.-Y., Montenay, J.-Y., Poitreau, J., Poucet, B., & Chaillan, F. A. (2020). Dorsal, but not ventral, hippocampal inactivation alters deliberation in rats. *Behavioural Brain Research*, *390*, 112622. https://doi.org/10.1016/j.bbr.2020.112622

Milner, B. (1954). Intellectual function of the temporal lobes. *Psychological Bulletin*, *51*(1), 42–62. https://doi.org/10.1037/h0054728

Milner, B. (1972). Disorders of Learning and Memory after Temporal Lobe Lesions in Man: Chapter 25. *Neurosurgery*, *19*, 421. https://doi.org/10.1093/neurosurgery/19.CN_suppl_1.421

Milner, B., & Penfield, W. (1955). The effect of hippocampal lesions on recent memory. *Transactions of the American Neurological Association*, *80th Meeting*, 42–48.

Morris, R. G. M., Garrud, P., Rawlins, J. N. P., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, *297*(5868), Article 5868. https://doi.org/10.1038/297681a0

Muenzinger, K. F. (1938). Vicarious Trial and Error at a Point of Choice: I. A General Survey of its

Relation to Learning Efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology*, *53*(1), 75–86. https://doi.org/10.1080/08856559.1938.10533799

Muenzinger, K. F., & Gentry, E. (1931). Tone discrimination in white rats. *Journal of Comparative Psychology*, *12*(2), 195–206. https://doi.org/10.1037/h0072238

Nádasdy, Z., Hirase, H., Czurkó, A., Csicsvari, J., & Buzsáki, G. (1999). Replay and Time Compression of Recurring Spike Sequences in the Hippocampus. *Journal of Neuroscience*, *19*(21), 9497–9507. https://doi.org/10.1523/JNEUROSCI.19-21-09497.1999

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. https://doi.org/10.1016/0006-8993(71)90358-1

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarendon Press ; Oxford University Press.

O'Keefe, J., & Nadel, L. (1979). Précis of O'Keefe & Nadel's *The hippocampus as a cognitive map*. *Behavioral and Brain Sciences*, *2*(4), 487–494. https://doi.org/10.1017/S0140525X00063949

Olton, D. S., Walker, J. A., & Gage, F. H. (1978). Hippocampal connections and spatial discrimination. *Brain Research*, *139*(2), 295–308. https://doi.org/10.1016/0006-8993(78)90930-7

Papale, A. E., Stott, J. J., Powell, N. J., Regier, P. S., & Redish, A. D. (2012). Interactions between deliberation and delay-discounting in rats. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(3), 513–526. https://doi.org/10.3758/s13415-012-0097-7

Papale, A. E., Zielinski, M. C., Frank, L. M., Jadhav, S. P., & Redish, A. D. (2016). Interplay between Hippocampal Sharp-Wave-Ripple Events and Vicarious Trial and Error Behaviors in Decision Making. *Neuron*, *92*(5), 975–982. https://doi.org/10.1016/j.neuron.2016.10.028

Pavlides, C., & Winson, J. (1989). Influences of hippocampal place cell firing in the awake state on the activity of these cells during subsequent sleep episodes. *Journal of Neuroscience*, *9*(8), 2907–2918. https://doi.org/10.1523/JNEUROSCI.09-08-02907.1989

Ranck, J. B. (1973). Studies on single neurons in dorsal hippocampal formation and septum in unrestrained rats: Part I. Behavioral correlates and firing repertoires. *Experimental Neurology*, *41*(2), 462–531. https://doi.org/10.1016/0014-4886(73)90290-2

Redish, A. D. (2016). Vicarious trial and error. *Nature Reviews Neuroscience*, *17*(3), 147–159. https://doi.org/10.1038/nrn.2015.30

Schmidt, B., Duin, A. A., & Redish, A. D. (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. *Journal of Neurophysiology*, *121*(6), 1981–2000. https://doi.org/10.1152/jn.00793.2018

Schmidt, B., Papale, A., Redish, A. D., & Markus, E. J. (2013). Conflict between place and response navigation strategies: Effects on vicarious trial and error (VTE) behaviors. *Learning & Memory*, *20*(3), 130–138. https://doi.org/10.1101/lm.028753.112

Scoville, W. B., & Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*(1), 11–21.

Seamans, J. K., Floresco, S. B., & Phillips, A. G. (1995). Functional differences between the prelimbic and anterior cingulate regions of the rat prefrontal cortex. *Behavioral Neuroscience*, *109*(6), 1063–1073. https://doi.org/10.1037/0735-7044.109.6.1063

Sigurdsson, T., Stark, K. L., Karayiorgou, M., Gogos, J. A., & Gordon, J. A. (2010). Impaired hippocampal–prefrontal synchrony in a genetic mouse model of schizophrenia. *Nature*, *464*(7289), Article 7289. https://doi.org/10.1038/nature08855

Skaggs, W. E., & McNaughton, B. L. (1996). Replay of Neuronal Firing Sequences in Rat Hippocampus During Sleep Following Spatial Experience. *Science*, *271*(5257), 1870–1873.

Stout, J. J., George, A. E., Kim, S., Hallock, H. L., & Griffin, A. L. (2023). *Harnessing prefrontal-hippocampal theta synchrony to enhance memory-guided choice* (p. 2023.04.02.535279). bioRxiv. https://doi.org/10.1101/2023.04.02.535279

Stout, J. J., Hallock, H. L., Adiraju, S. S., & Griffin, A. L. (2021). *The Ventral Midline Thalamus Mediates Successful Deliberation by Coordinating Prefrontal and Hippocampal Neural Activity* (p. 2021.06.30.450519). https://doi.org/10.1101/2021.06.30.450519

Tang, W., Shin, J. D., & Jadhav, S. P. (2021). Multiple time-scales of decision-making in the hippocampus and prefrontal cortex. *ELife*, *10*, e66227. https://doi.org/10.7554/eLife.66227

Tavares, L. C. S., & Tort, A. B. L. (2022). Hippocampal–prefrontal interactions during spatial decision-making. *Hippocampus*, *32*(1), 38–54. https://doi.org/10.1002/hipo.23394

Tolman, E. C. (1926). A behavioristic theory of ideas. *Psychological Review*, *33*(5), 352–369. https://doi.org/10.1037/h0070532

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. https://doi.org/10.1037/h0061626

Tolman, E. C., Ritchie, B. F., & Kalish, D. (1946). Studies in spatial learning II Place learning versus response learning. *Journal of Experimental Psychology*, *36*(3), 221–229. https://doi.org/10.1037/h0060262

Vinogradova, O. S. (1970). Registration of information and the limbic system. In *Short-term changes in neural activity and behavior* (pp. 95–140). Cambridge: University Press.

Wang, G.-W., & Cai, J.-X. (2006). Disconnection of the hippocampal–prefrontal cortical circuits impairs spatial working memory performance in rats. *Behavioural Brain Research*, *175*(2), 329–336. https://doi.org/10.1016/j.bbr.2006.09.002

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of Hippocampal Ensemble Memories During Sleep. *Science*, *265*(5172), 676–679.

Wood, E. R., Dudchenko, P. A., Robitsek, R. J., & Eichenbaum, H. (2000). Hippocampal Neurons Encode Information about Different Types of Memory Episodes Occurring in the Same Location. *Neuron*, *27*(3), 623–633. https://doi.org/10.1016/S0896-6273(00)00071-4

Yerkes, R. M. (1907). *The dancing mouse: A study in animal behavior*. MacMillan Co. https://doi.org/10.1037/10935-000

Zielinski, M. C., Shin, J. D., & Jadhav, S. P. (2019). Coherent Coding of Spatial Position Mediated by Theta Oscillations in the Hippocampus and Prefrontal Cortex. *Journal of Neuroscience*, *39*(23), 4550–4565. https://doi.org/10.1523/JNEUROSCI.0106-19.2019

Zola-Morgan, S., Squire, L. R., & Amaral, D. G. (1986). Human amnesia and the medial temporal region: Enduring memory impairment following a bilateral lesion limited to field CA1 of the hippocampus. *Journal of Neuroscience*, *6*(10), 2950–2967. https://doi.org/10.1523/JNEUROSCI.06-10-02950.1986

Chapter 2

A machine learning approach for detecting vicarious trial and error behaviors

A version of this chapter has been published in the journal *Frontiers in Neuroscience*.

Citation:
Miles JT, Kidder KS, Wang Z, Zhu Y, Gire DH, and Mizumori SJY (2021). A Machine Learning Approach for Detecting Vicarious Trial and Error Behaviors. Front. Neurosci. Jul. 2021.
doi: 10.3389/fnins.2021.676779

Abstract

Vicarious trial and error behaviors (VTEs) indicate periods of indecision during decision-making, and have been proposed as a behavioral marker of deliberation. In order to understand the neural underpinnings of these putative bridges between behavior and neural dynamics, researchers need the ability to readily distinguish VTEs from non-VTEs. Here we utilize a small set of trajectory-based features and standard machine learning classifiers to identify VTEs from non-VTEs for rats performing a spatial delayed alternation task (SDA) on an elevated plus maze. We also show that previously reported features of the hippocampal field potential oscillation can be used in the same types of classifiers to separate VTEs from non-VTEs with above chance performance. However, we caution that the modest classifier success using hippocampal population dynamics does not identify many trials where VTEs occur, and show that combining oscillation-based features with trajectory-based features does not improve classifier performance compared to trajectory-based features alone. Overall, we propose a standard set of features useful for trajectory-based VTE classification in binary decision tasks, and support previous suggestions that VTEs are supported by a network including, but likely extending beyond, the hippocampus.

2.1 Background and Introduction

Introduced and popularized in the 1930s, vicarious trial and error (VTE) is a well-documented behavioral phenomenon where subjects vacillate between reward options before settling on their final choice (Muenzinger and Gentry, 1931; Tolman, 1938). This behavior is best understood in rats making decisions to go left or right, and as such, VTE trajectories tend to have curves that change direction at decision points. Current theories claim that subjects mentally assess possible options before making a final decision during VTEs (Redish, 2016), suggesting that they may be related to, but not necessarily identical to, an underlying deliberative process. While such a relationship to deliberation is complex and outside of the scope of this paper, it is clear that VTEs are a valuable behavioral variable to take into account when studying decision-making, particularly when investigating neural processing during decisions.

The majority of recent experiments examining the neural underpinnings of VTEs have focused on the rodent hippocampus (HPC). Bilateral electrolytic HPC lesions decrease the mean number of VTEs in a visual discrimination task, particularly during early learning (Hu and Amsel, 1995); though see Bett et al. (2012). Similarly, bilateral ibotenic acid HPC lesions decrease the number of VTEs rats exhibit before they have located a reward in a spatial task (Bett et al., 2012). In addition to lesion studies, several electrophysiological findings link the HPC to VTEs. Dorsal HPC recordings during VTEs show serial sweeps of place cell sequences, which first trace the initial direction of the VTE before sweeping in the direction a rat ends up choosing (Johnson and Redish, 2007). Furthermore, dorsal HPC place cell recordings are more likely to represent locations of an unchosen option during VTEs than non-VTEs (Papale et al., 2016). There is also evidence that the field potential oscillation recorded from dorsal HPC differs on decisions where VTEs do and do not occur. In particular, characteristics of HPC theta (4–12 Hz) oscillations, such as its shape and duration, appear to be altered during VTEs, as do aspects of gamma-band (35–100 Hz) oscillations (Amemiya and Redish, 2018; Schmidt et al., 2019), but see Dvorak et al. (2018).

Despite decades-long interest and their utility as a behavioral marker of a putative cognitive process, VTEs have been studied by only a small number of labs. We suspect part of the reason they have not received more attention is that VTE trajectories can be highly variable, which makes it difficult to identify them algorithmically (Goss and Wischner, 1956). The Redish lab has proposed the zIdPhi metric, which quantifies changes in heading angles as rats traverse choice points, for identifying VTEs (Papale et al., 2012, 2016; Amemiya and Redish, 2016, 2018; Redish, 2016; Schmidt et al., 2019; Hasz and Redish, 2020). While successful in their hands, zIdPhi, admittedly, “does not provide a sharp boundary between VTE and not” (Papale et al., 2016, Supplementary Material, section Experimental Procedures). Here we show that standard machine learning models trained on data from a spatial delayed alternation task are able to robustly and reliably distinguish VTE trajectories from non-VTEs in such a task.

Additionally, we assess how the same types of classifier models perform when trained on features of the dorsal HPC oscillation that have been shown to differ between VTEs and non-VTEs (e.g., differences in gamma power and theta wave shape; Amemiya and Redish, 2018; Schmidt et al., 2019). In doing so, we demonstrate that these features are indeed able to separate decision types better than would be expected by random binary classification, though with worse performance than trajectory-based features. Furthermore, we show that providing a classifier with HPC oscillatory dynamics from when animals make choices yields better performance than oscillations from the immediately preceding delay interval, which is when information about the previous choice would need to be held in memory. We also show that a more comprehensive description of the HPC oscillation, the power spectrum, does not perform any better than the model trained on curated features. Finally, we demonstrate that combining informative trajectory- and oscillation-based features does not change classifier performance when compared to classification using trajectory features alone, leading us to conclude that the HPC oscillation does not contain information that complements what can be extracted from the trajectories.

2.2 Results

2.2.1 Trajectory-based classification

A VTE occurs when rats vacillate between options before their final choice. Behaviorally, this manifests as a trajectory with curves or sharp angles at choice points, where reorientations occur (**Figure 2.1A**). We analyzed a dataset with 828 trajectories from rats running a SDA task (Baker et al., 2019; Kidder et al., 2021). Each trajectory was scored as VTE (n = 142) or non-VTE (n = 686) by four trained annotators. We calculated zIdPhi (Papale et al., 2012), the z-scored, integrated change in heading angle, for each trajectory, as well as several other features (see 2.4.1 for more details). As expected, we saw statistically distinct empirical distributions for zIdPhi values on trials scored as VTE compared to non-VTE (**Figure 2.1B**, $p < 0.001$, two-sample K-S test; $d = 0.68$). When compared to manual scoring, however, using zIdPhi did not reliably separate VTE and non-VTE trials (**Figure 2.2A**).

We reasoned that we could obtain more accurate and reliable VTE detection by assessing multiple aspects of the trajectory instead of just one. As such, we calculated seven trajectory-based features (see Classifier Implementation for details) with the expectation that these features would allow for separation of VTEs and non-VTEs in a higher dimensional space. Like zIdPhi, many of these

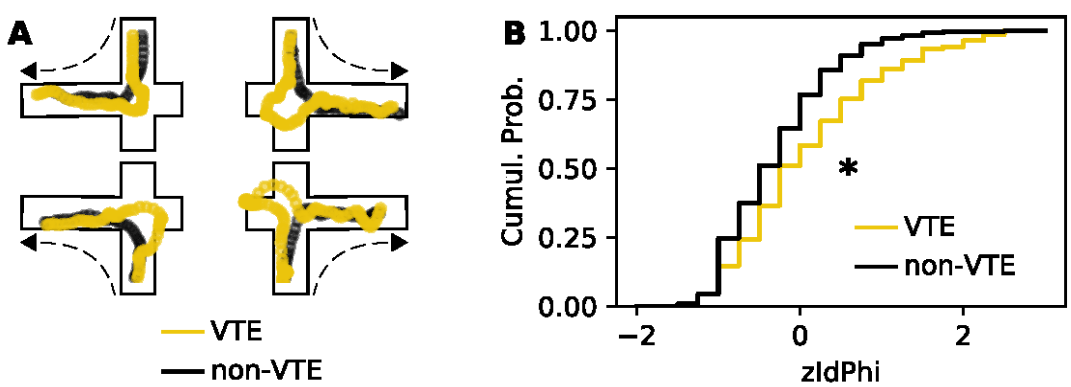

Figure 2.1: *Example VTEs and zIdPhi distribution.* **(A)** Example trajectories showing VTEs (yellow) and non-VTEs (black). Trajectories are shown on top of outlines of the decision-point, coming from each possible direction (shown by dashed arrow). **(B)** Empirical cumulative distributions of zIdPhi scores for VTEs (yellow) and non-VTEs (black). Note the prominent rightward shift for VTEs ($p < 0.001$, two-sample K-S test, $d = 0.65$). *Cumul. Prob., cumulative probability; K-S, Kolmogorov-Smirnov.*

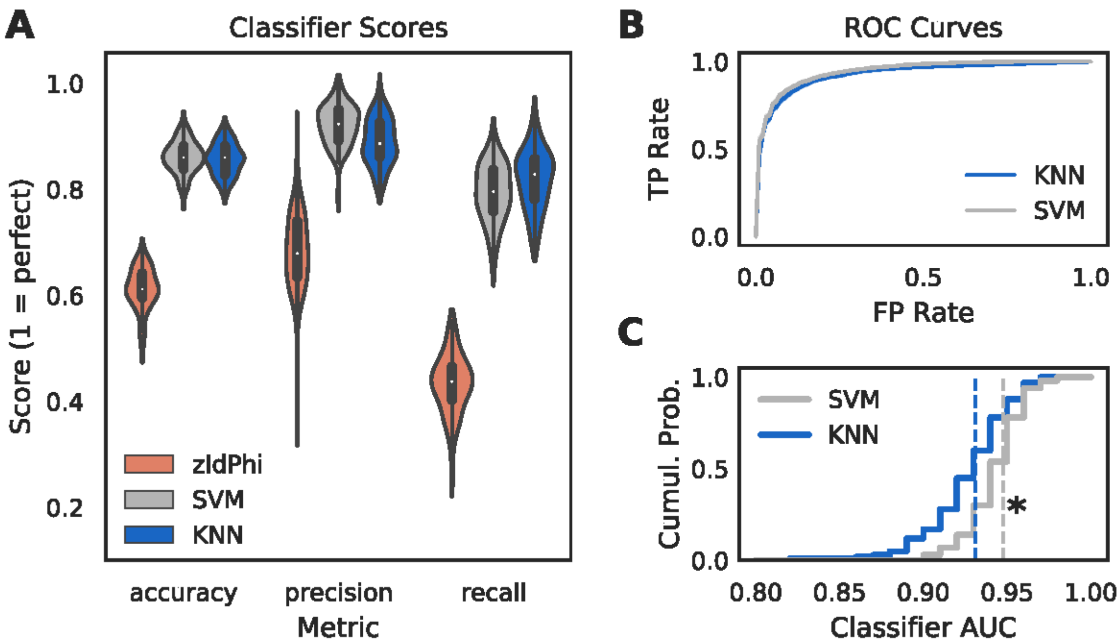

Figure 2.2: *Multi-feature classifiers outperform a single metric.* **(A)** Classification performance on 100 random splits of manually scored data using zIdPhi (orange), an SVM model (gray), and a KNN model (blue). Colored patches are kernel density estimates of the underlying distributions, with boxplots representing the same data inside the patches. **(B)** Receiver operating characteristic curves for KNN (blue) and SVM (gray) classifiers. Averages across splits are bold lines, 95% confidence intervals surround means, but are not visible. **(C)** Cumulative distributions of AUC scores for SVM in gray and KNN in blue. Note the rightward shift for the SVM model ($p < 0.0001$, two-sample K-S test, $d = 0.78$). *SVM, support vector machine; KNN, k-nearest neighbor; AUC, area under the curve; TP, true positive; FP, false positive; Cumul. Prob., cumulative probability*

features formed distinct empirical distributions for VTEs and non-VTEs, which suggested to us that this feature set could beused to build machine learning classifiers for algorithmic VTE detection.

Classifiers are often evaluated for their accuracy, precision, and recall scores (Malley et al., 2011; Lever et al., 2016) (see **Classifier implementation** for detailed descriptions). In the context of VTE identification, accuracy measures the proportion of correctly labeled trials (i.e. VTE or non-VTE), precision measures the proportion of trials labeled VTE that are actually VTEs, and recall measures the proportion of VTEs found out of the total number of VTEs present. We compared performance of two widely used machine learning models – k-nearest neighbors (KNN) and support vector machines (SVM) – to zIdPhi alone in **Figure 2.2**. To generate distributions for each of these metrics, we scored 100 iterations of randomly sampled splits of data, with mutually exclusive testing and training trajectories (see **Classifier implementation** for further details). To ensure scores were not influenced by the fact that we had many more non-VTE trials than VTE trials, we equalized the number of VTE and non-VTE trials for each data split. Both KNN and SVM classifiers show high accuracy ($\overline{A}_{knn} = 0.86$, $\overline{A}_{svm} = 0.86$; bars above letters denote mean), precision ($\overline{P}_{knn} = 0.86$, $\overline{P}_{svm} = 0.88$), and recall ($\overline{R}_{knn} = 0.80$, $\overline{R}_{svm} = 0.82$) on our trajectory data. Comparing the different classifiers' distributions of a composite precision and recall score, the F_1

score, shows that the performance for the SVM classifier is generally higher ($\overline{F1}_{knn}$ = 0.82, $\overline{F1}_{svm}$ = 0.85; p = 0.0001, two-sample K-S test). Overall, these results suggest that we have defined a feature set suitable for VTE classification, that both KNN and SVM models provide more accurate, sensitive, and precise VTE classification than a single metric alone, and that the SVM model has a slight performance edge over the KNN model.

2.2.2 Oscillation-based classification

Previous research has suggested HPC involvement in decisions where VTEs occur. Early work showed that rats with bilateral HPC lesions perform less VTEs during initial learning in a visual discrimination task than rats with their hippocampi intact (Hu and Amsel, 1995). More recent research did not find differences in VTE rates for lesioned and non-lesioned animals during visual discrimination, but showed that lesioned rats exhibit fewer VTEs during early learning when performing a spatial decision-making task. In particular, lesioned rats showed fewer VTEs before finding a new reward location after it had been moved (Bett et al., 2012). Additionally, multiple studies have shown that HPC place cell activity is more likely to represent future locations during decisions involving a VTE than when no VTE occurs (Johnson and Redish, 2007; Papale et al., 2016). Furthermore, several features of the hippocampal local field potential oscillation appear to be different when decisions are made with, as opposed to without, VTEs (Amemiya and Redish, 2018; Schmidt et al., 2019).

We tested how well-features of the HPC oscillation (**Figure 2.3A**) could identify VTEs using the same approach we employed for trajectory-based VTE classification. Consistent with previous work, we found several oscillatory features with different empirical distributions for VTE and non-VTE trials (**Figure 2.3**). To test whether an SVM classifier could identify VTEs above chance levels when trained with features of the HPC oscillation, we calculated classifier metric Δ scores.

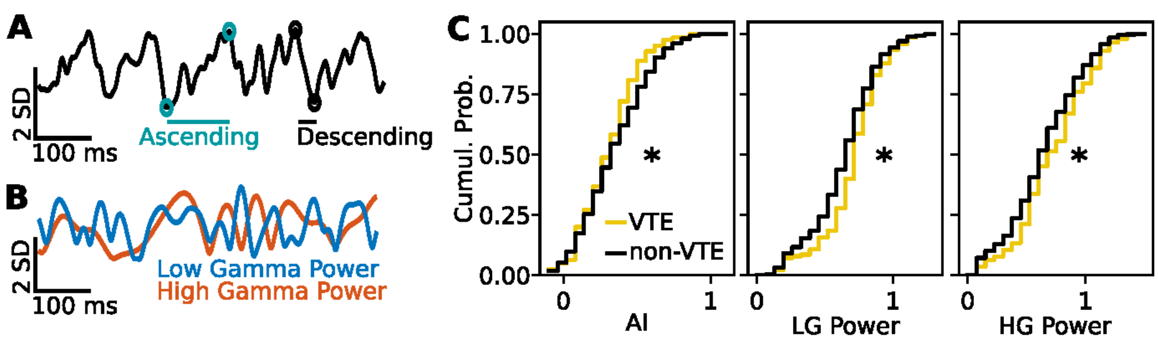

Figure 2.3: *Example oscillation data and feature distributions.* **(A)** Sample HPC oscillation. Data are normalized (z-scored), so amplitude is measured in standard deviations (see scale bar). Blue-green circles and line show the ascending duration of one theta cycle, and black circles and line show the descending duration of the next theta cycle. The ascending and descending durations within a single theta cycle are used to calculate the AI. **(B)** Normalized power timeseries for low gamma (blue) and high gamma (orange) for the oscillation shown in **(A)**. **(C)** Cumulative distributions for 3 (of 12) curated features of the HPC oscillation. Asterisks denote significantly different distributions between VTE and non-VTE trials (AI, p = 0.041; LG Power, p < 0.001; HG Power p = 0.049).

AI, asymmetry index; SD, standard deviation; ms, millisecond; LG, low gamma; HG, high gamma.

We compared classifier performance on hippocampal data from two distinct behavioral epochs—one where rats actively made choices (i.e., when VTEs would occur), or during the delay interval that preceded the choice epoch. Each score shows how far above chance the classifier performed when oscillations were taken from the choice or delay epoch (**Figure 2.4A**). Chance estimates were obtained by training a classifier on oscillations from the choice epoch, but randomly labeling each trial as VTE or non-VTE. Thus, a score of zero indicates that the classifier performed the same as would be expected if randomly labeling trials. Classifier performance on the HPC oscillation during choices is above the performance for classifiers trained on the HPC oscillation during the delay epoch of the task (**Figure 2.4B**; $\overline{\Delta AUC}_{delay} = 0.004$, $\overline{\Delta AUC}_{choice} = 0.11$; two-sample K-S test, $p < 0.0001$; $d = 1.21$).

Though the highly curated features used in the classifier for **Figure 2.4** have been shown to differ during VTEs and non-VTEs (Amemiya and Redish, 2018; Schmidt et al., 2019), these features are only a small subset of attributes that could describe the HPC field potential oscillation. As such, we used arguably the most common descriptor of oscillations, the power spectrum, in an attempt to increase classifier performance. We first compared average power spectral density (PSD) estimates for different frequencies, calculated for different splits of data, to identify which frequencies had significantly different average power on VTE and non-VTE decisions. Frequencies that survived false discovery rate correction were used as features for an

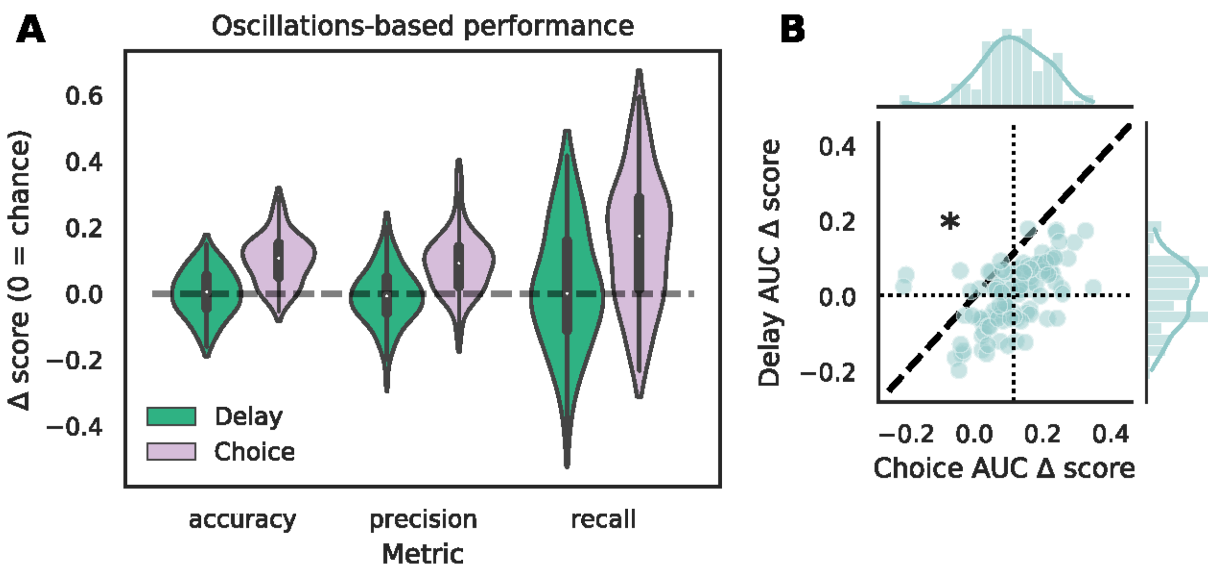

Figure 2.4: *VTE classification is better for oscillation-based features from choices than delays.* **(A)** Classification performance using oscillations from the delay epoch (green) vs oscillations from the choice epoch (purple). Each Δ score is a difference between performance using data from either epoch and randomly labeled data. **(B)** Scatterplots of ΔAUC score distributions for oscillations taken from the delay and choice epochs. Dotted vertical line shows the mean choice ΔAUC score, dotted horizontal line shows the mean delay ΔAUC score. Dashed diagonal line marks where equal departures from chance would occur. Individual distributions and kernel density estimates for ΔAUC scores are shown in the marginal distributions. Note skew below the diagonal, indicating significantly higher choice ΔAUC scores ($p < 0.0001$, two-sample K-S test, $d = 1.21$).

SVM classifier trained on PSD estimates. Interestingly, although these classifiers utilized a much higher dimensional feature-space (roughly seven-fold more features using PSD estimates than curated oscillation-based features), ΔAUC scores were no different from those obtained with the highly curated features (**Figure 2.5B**, $p = 0.70$, two-sample K-S test; $d = 0.20$).

It is possible that features of the HPC oscillation contain information about VTE occurrence that complements the information contained in trajectory data. In other words, VTEs that are difficult to classify based on trajectories alone may have accompanying HPC oscillatory dynamics that, when combined with the trajectory features, lead to improved VTE classification. To examine this possibility, we trained an SVM classifier on combined trajectory- and oscillation-based features from the choice epoch that had significantly different distributions on VTE and non-VTE trials. Interestingly, combining feature sets does not change performance when compared to trajectory features alone (**Figure 2.6**; mean AUC difference = −0.001; $p = 0.65$, Wilcoxon signed rank test; $d = 0.08$). Thus, we conclude that, although features of the HPC oscillation can be used to some extent for classifying VTEs, these features do not contain novel or complementary information beyond what can be extracted from the trajectories.

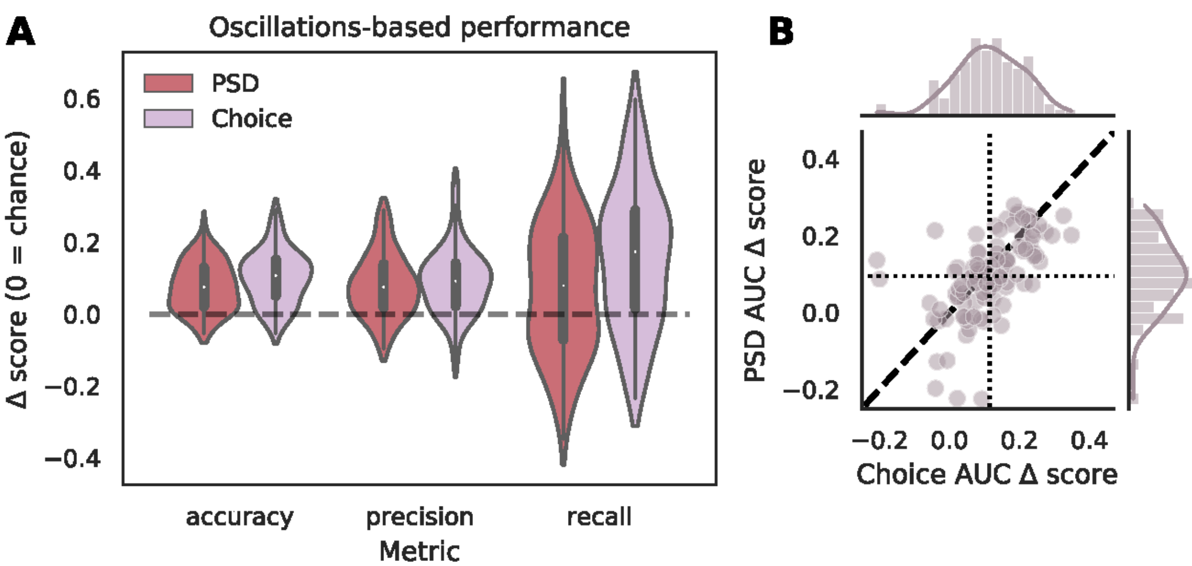

Figure 2.5: *Full power spectra do not outperform highly curated oscillation features*. **(A)** Classification performance when using PSD estimates (dark red) compared to curated features from the choice epoch (purple). Each Δ score is measured as a difference between performance using either PSD or oscillation-based features and corresponding randomly labeled data. **(B)** Scatterplots of ΔAUC score distributions for PSD estimates and oscillation-based features, both calculated for choice epochs. Dotted vertical line shows the mean choice oscillation-based ΔAUC score, dotted horizontal line shows the mean PSD ΔAUC score. Dashed diagonal line marks where equal departures from chance would occur. Individual distributions and kernel density estimates for ΔAUC scores are shown in the marginal plots. ΔAUC scores are not significantly different for SVMs trained on PSD-based and curated feature sets ($p = 0.70$, two-sample K-S test, $d = 0.20$). *PSD, power spectral density.*

2.3 Discussion

The purpose of this study was to improve upon current methods of VTE identification and build on our understanding of hippocampal involvement during VTEs. We show that VTE behavior can be robustly and reliably separated from non-VTE behavior using a small set of trajectory-based features. Additionally, we show that classifiers trained on features of the dorsal HPC field potential oscillation separate VTEs from non-VTEs more than would be expected by chance, supporting previous research linking the HPC to VTEs. Moreover, we show that when oscillations are taken from the delay epoch that precedes the choice epoch, oscillation-based features no longer enable above chance VTE classification, which suggests a brief temporal window underlying HPC involvement in VTE processing. We also caution, however, that despite above chance VTE identification using oscillation-based features, our results also clearly show that population level HPC dynamics are prone to VTE misclassification (**Figures 1.4, 1.5**), especially when compared to trajectory-based features. In particular we demonstrate that combining neural features and trajectory features does not improve performance compared to using trajectory features alone.

Not only do the small set of hippocampal oscillation features previously reported to differ between decisions where VTEs do and do not occur provide modest performance for VTE classification, using the roughly seven-fold larger feature space of the 1–100 Hz hippocampal power spectrum does not improve performance. We see this as further evidence that population level HPC dynamics only partially explain VTEs. We suspect that examining HPC interactions with other areas, such as the mPFC (Brown et al., 2016; Voss and Cohen, 2017; Schmidt et al.,

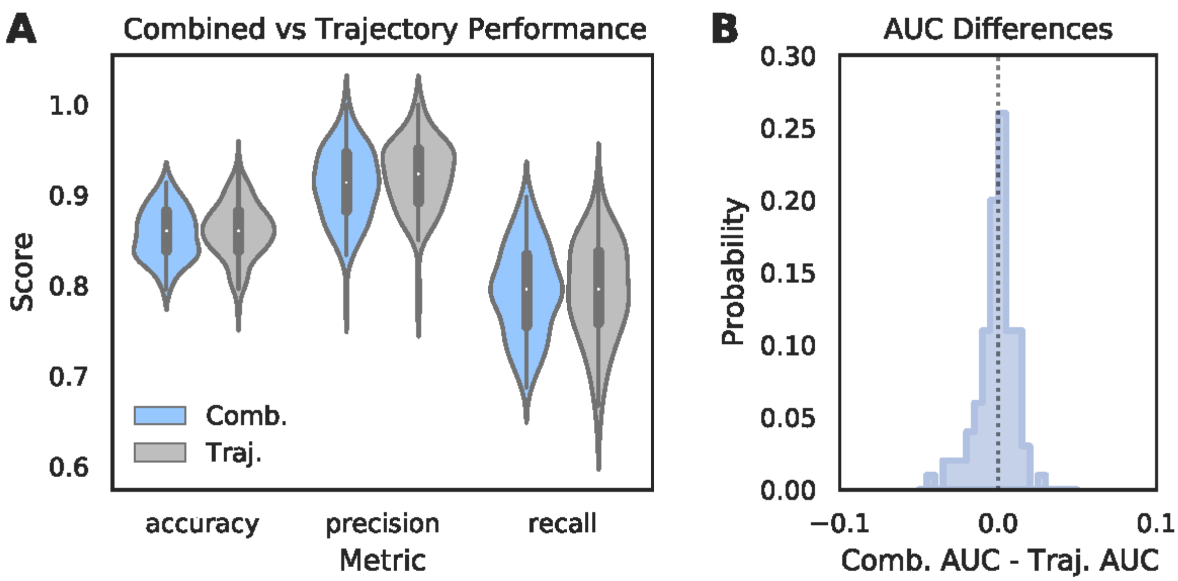

Figure 2.6: *Combining oscillation- and trajectory- based features does not increase classifier performance.* **(A)** Comparison of SVM classifier performance using combined oscillation- and trajectory- based features (blue) or trajectory features alone (gray, same as **Figure 2.2A**—SVM). **(B)** Distribution of *AUC* scores for classifiers trained on trajectory only features subtracted from *AUC* scores for classifiers trained on combined feature sets for corresponding data splits. Values below zero indicate worse performance using combined features. Scores are centered around zero (p = 0.65, Wilcoxon rank sum test, d = 0.08). *Comb., combined; Traj., trajectory alone.*

2019; Hasz and Redish, 2020; Kidder et al., 2021), would be a fruitful next step for improving our ability to classify VTEs based on neural activity. Schmidt et al. have shown that rats perform fewer VTEs when faced with difficult decisions if their mPFC has been inhibited chemogenetically. Furthermore, the window in which the HPC oscillation is best able to identify VTEs is during choices, which is when brief increases in theta coherence between the dorsal HPC and mPFC occur (Jones and Wilson, 2005; Benchenane et al., 2010), suggestive of cross-regional communication (Fries, 2005, 2015). Finally, experiments using optogenetics to perturb the mPFC in a task-epoch-specific manner during the SDA task showed that stimulation decreased the proportion of VTEs rats engaged in, with a trend toward choice epoch mPFC disruption having a greater effect than stimulation in other epochs (Kidder et al., 2021).

Methodologically, we find comparing classification performance between behavior and neural activity an intuitive way to understand how well the activity under scrutiny relates to the behavior in question. The level of performance for behavior classification can often be thought of as an upper bound for assessing how well neural activity describes the behavior, while randomly labeled classifiers can set the lower bound. This may provide a more nuanced picture of how well neural activity relates to a behavior than hypothesis testing alone. For example, while we and others show multiple features of the HPC oscillation form distinct empirical distributions for VTEs and non-VTEs, the fact that classifier performance using these features does not meet classification performance of the behavior itself suggests that these features only provide a partial description about the neural substrate of the behavior. Additionally, feature-based classification allows for very flexible control of what parameters—behavioral or neural—one wishes to examine, as well as the size of the parameter space one would like to search. Moreover, as demonstrated by comparing HPC power spectra with curated oscillation features, feature vectors can be arbitrarily sized with surprisingly little influence on classifier performance, as long as the classifier is constructed to protect against overfitting (e.g., with proper hyper-parameter selection and cross-validation). For these reasons, we see this framework as extremely flexible in terms of feature selection and use, as well as an intuitive way of gauging how well neural activity measurements describe behavior.

A limitation of our study is that we do not explicitly test which features of the hippocampal oscillation are the best indicators of VTE behavior, nor do we claim that the features we test are an exhaustive list of possible features. Rather, we ask if oscillation-based features suggested by prior work can identify VTE behaviors, and to what extent they match the ability of a classifier using trajectory-based features. Similarly, this study does not address whether there is an optimal subset of power spectral density features for VTE identification. Instead, we specifically ask to what extent the range of frequencies from 1 to 100 Hz is able to identify VTE behaviors. Thus, we leave open the possibility that the hippocampal oscillation may be better able to explain VTE behaviors than is reported in this study, while suggesting a framework that others can build on to test their own hypotheses.

Altogether, our results expand previous efforts to algorithmically identify VTEs using choice trajectories from a given behavioral task, improving our ability to detect these important variants of decision-making behavior. In addition, we provide further evidence for hypotheses that situate the hippocampus as one element in what is likely a broader network of interacting neural structures supporting VTEs. We believe future decision-making research will benefit from tracking VTEs and VTE-like behaviors, such as saccades and head movements in humans and non-human primates (Voss and Cohen, 2017; Santos-Pata and Verschure, 2018) and hope our classification scheme enables more wide-spread VTE analysis. Additionally, we encourage future VTE research to expand beyond the HPC and further our understanding of the neural system(s) involved in this decision-making behavior.

2.4 Methods

2.4.1 Behavioral task

Food restricted (85% of body weight) Long Evans rats (n = 9, Charles River Laboratories) were trained on a previously described spatial delayed alternation (SDA) task (Baker et al., 2019), (Kidder et al., 2021). Briefly, sessions were run on an elevated plus maze (black plexiglass arms, 58 cm long x 5.5 cm wide, elevated 80 cm from floor), with moveable arms and reward feeders controlled by custom LabView 2016 software (National Instruments, Austin, TX, USA). Each trial consisted of a rat leaving its starting location in a randomly chosen "north" or "south" arm, then navigating to an "east" or "west" arm for a 45 mg sucrose pellet reward (TestDiet, Richmond, IN, USA). Rewards were delivered when rats alternated from their previously chosen arm (i.e. if they selected the "east" arm on trial $n - 1$ then they had to select the "west" arm for reward on trial n). After making a choice, rats had the opportunity for reward consumption (if correct) before they returned to the assigned start arm and entered into a 10 second delay period before the next trial began. Based on this structure, we divided the task into three epochs - choice, return, and delay.

2.4.2 Microdrive implantation

Micro-drive bodies were 3-D printed (Form 2 Printer; Formlabs, Sommerville, MA) to contain between 8 and 16 gold plated tetrodes (nichrome; SANDVIK, Sandviken, Sweden), which were implanted unilaterally into the CA1 region of HPC (AP: -3.0, M/L: ±2.0mm, D/V: -1.8mm). A subset of animals (3) had two optic fibers that were implanted bilaterally into the mPFC and used for additional experiments (Kidder et al., 2021), and the remainder (6) also had tetrodes implanted into the ipsilateral lateral habenula for additional experiments, but all animals ran the same behavioral task, and data used for this study were from before any optogenetic stimulation was ever delivered. Tetrodes were connected to a 64-channel Open Ephys electrode interface board (EIB) (open-ephys.org). To eliminate external noise, drive bodies were shelled in plastic tubes lined with aluminum foil coated in a super-conductive nickel spray. One ground wire connected the shell with the EIB and, during surgery, another ground wire was implanted near the cerebellum just inside the skull. After surgery, rats were allowed to recover for approximately seven days before entering into testing, and HPC tetrodes were lowered over the course of several days until at least one tetrode showed oscillations consistent with the CA1 fissure (high-amplitude, asymmetric theta oscillations).

2.4.3 Data Acquisition

Behavior tracking

Two LEDs were attached to either the rat's microdrive or the tethers plugged into the microdrive's head- stage before recordings. Rat locations were determined by subtracting a background image taken at the beginning of the session from each frame. Pixels containing the LEDs showed an above threshold difference in brightness, which allowed us to determine rat head locations in each frame. Camera frames were recorded at approximately 35 Hz using a SONY USB web camera (Sony Corporation, Minato, Tokyo). Frames were time-stamped with a millisecond timer run by LabView and sent to the Open Ephys acquisition software (open-ephys.org) for later alignment of electrophysiological and position information.

Electrophysiology

Electrophysiological data were sampled at 30 kHz using Intan headstages (RHD2132; Intan Technologies, Los Angeles, California) connected to the Open-Ephys EIB. Digitized signals were sent via daisy chained SPI cables through a motorized commutator that prevented tether twisting (AlphaComm-I; Alpha Omega Co., Alpharetta, GA) and into an Open-Ephys acquisition board (open-ephys.org). All further processing and filtering was done offline using custom MATLAB scripts (see *Oscillation-Based Features* for more details).

2.4.4 Classifier features - trajectories

We calculated 7 features of choice epoch trajectories - the standard deviation (SD) of the x-position (x_σ), the SD of the y-position (y_σ), the trial's z-scored, integrated change in heading angle (*zIdPhi*), the trial duration (*dur*), how well the trial was fit by a sixth degree polynomial (r^2), and the number of Fourier coefficients needed to describe the fit of the polynomial (n_{coef}). Both x_σ and y_σ were calculated using the std method from `Python's numpy` package for the x and y position vectors of the rat's trajectory on a given trial. The *IdPhi* score for a trial was defined as:

$$\varphi = arctan2(dy, dx) \tag{1}$$

$$IdPhi = \sum_{i=a}^{n} |\varphi_a - \varphi_{a-1}| \tag{2}$$

where *arctan*2 is the 2-argument arctangent function and *dx* and *dy* are changes in the trajectory's x and y position, respectively. This value was transformed to *zIdPhi* by converting to a z-score, which was calculated for each session individually. We set the *zIdPhi* threshold value, above which something was assigned as a VTE, by iterating through values from the 50th to the 80th percentile and choosing the value that maximized classification accuracy. The *dur* feature measures the duration a rat was within an experimenter defined choice point on the maze. The r^2 value was determined using a two-step process. First, optimal coefficients for each of the terms in the polynomial were calculated using the `curve_fit` method of the `scipy.optimize` package with the vector of x positions as the independent variable and the vector of y positions as the dependent variable. From here, we used the optimized outputs as inputs to a generic sixth degree polynomial function, calculated the error sum of squares between the observed y values and modeled outputs (*SSE*, see equation 3), and calculated the total sum of squares (SST, see equation 4). The calculation of the r^2 value is shown in equation 5.

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3}$$

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{4}$$

$$r^2 = 1 - \frac{SSE}{SST} \tag{5}$$

In (3), $\hat{y}_i$ is the estimated y position at the i-th location in the trajectory, and $\bar{y}$ is the mean y position of the trajectory in (4). We noticed that plotting the polynomial estimates with poor fits (which were mainly VTEs) created a trajectory that looked similar to a damped oscillation, so we devised the n_{coef} feature - which is the number of Fourier coefficients needed to describe the polynomial fit estimate - to capture this oscillatory character. Intuitively, higher values of n_{coef} were expected to correlate with instances of VTE.

2.4.5 Classifier features

Curated features

To quantify features of the HPC CA1 oscillation, we down-sampled our data by a factor of 30, going from a 30kHz sampling rate to a 1kHz sampling rate, and z-scored the downsampled timeseries to put amplitude in units of standard deviations. Based on previous work (Amemiya and Redish, 2018; Schmidt et al., 2019), we were interested to see if we could use features of the HPC CA1 oscillation to classify VTE vs non-VTE trials. We used 7 features - the asymmetry index (*AI*) of the wide-band theta oscillation, average ascending (*asc*) and descending (*desc*) durations of the wide-band theta oscillation, the average (normalized) low and high gamma powers (G_L and G_H, respectively), the cycle-averaged gamma ratio (*GR*), and the average duration of a theta cycle. Each trial had multiple measurements of each value, so we also used the SD of these measurements as a feature for all but the *asc* and *desc* features, giving a 12-variable feature vector for each trial.

Previous reports have demonstrated asymmetric theta oscillations in different layers of the HPC (Buzsáki et al., 1985, 1986; Buzsáki, 2002), so we used a low-pass filtered signal with the cutoff frequency at 80 Hz to identify peaks and troughs of the theta oscillation as well as the ascending duration, descending duration, and total duration of each theta cycle (Belluscio et al., 2012) (see **Figure 2.3A** for an example). The *AI* is defined as:

$$AI = log(asc) - log(desc) \tag{6}$$

such that cycles with longer ascending than descending durations will give positive values, cycles with equal ascending and descending durations will equal 0, and cycles with shorter ascending than descending durations will give negative values. Because different HPC recording locations can have differently shaped theta oscillations (Buzsáki et al., 1985, 1986; Buzsáki, 2002), we ensured that all days used for analysis had *AI* distributions that were skewed in the same direction.

To estimate gamma powers, first we bandpass filtered our downsampled timeseries between 35-55 Hz for low gamma and 61 - 100 Hz for high gamma using third order, zero-lag Butterworth filters. These values were then z-scored, putting units of amplitude into standard deviations. The power in a gamma-band timeseries, g(t), was estimated using:

$$p(t) = |\tilde{g}(t)|^2 \tag{7}$$

where $\tilde{g}(t)$ denotes the Hilbert transform of $g(t)$. We then used these power estimates to calculate cycle-by- cycle *GRs*. For a given cycle, the gamma ratio was defined as:

$$GR = \frac{\hat{G}_L}{\hat{G}_H} \tag{8}$$

and the gamma ratio for the entire trial was the average of these cycle-by-cycle values.

Power spectral densities

In addition to pre-defined oscillation bands and bandpass filtering signals, we performed the same classifier- based analysis of neural data using PSD estimates as features instead of the curated oscillation features. For this, we used MATLAB's `periodogram` function (version 2018 B; MathWorks, Nattick, MA), with a Hamming window over the duration of the signal, a frequency resolution of 1 Hz, and a range of 1 Hz to 100 Hz. To maintain consistency with curated oscillation features, we use the z-transformed HPC oscillation. PSD estimates are kept as original values, as opposed to the common decibel conversion.

2.4.6 Classifier implementation

Classifier models

We used the `scikit_learn` library from `Python` to create and test k-nearest neighbor (KNN) and support vector machine (SVM) models. All instances of the KNN model used 5 neighbors for classification, though results for 3-10 neighbors did not lead to different conclusions. All instances of the SVM model used a radial basis function (RBF) kernel for assessing distance/similarity. A γ parameter dictates the width and shape of the RBF, with lower values giving wider kernel functions and higher values giving narrower kernel functions. We chose to search values between 0.005 and 10 for γ. Another parameter, the C parameter, controls the trade-off between the size of the decision function margin and classification accuracy, which can be thought of as a way to control overfitting the decision function. Low values of C favor a larger margin, high values of C favor a more complex decision function. We tested a range of C values from 0.01 to 10. Data used to train the model were standardized and scaled. Testing data given to the model were transformed based on the scalings calculated for the training data (see *Cross Validation* for more details on how data were used for classifier training and testing).

Evaluating classifiers

We used several standard metrics for assessing classifier performance (Lever et al., 2016), all of which de- scribe different combinations and/or weightings of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). For VTE identification, a TP is trial correctly classified as a VTE, a TN is a trial correctly classified as a non-VTE, a FP is a trial incorrectly classified as a VTE, and a FN is a trial incorrectly classified as a non-VTE. Accuracy measures the number of trials assigned to the correct class (VTE or non-VTE) out of the total number of trials, and is defined as:

$$accuracy = \frac{TP + TN}{TP + TN + FN + FP}$$

such that accuracy equals 1 if every trial, VTE and non-VTE, is correctly classified, and 0 if no trials are correctly classified. Precision measures the number of correctly classified VTEs out of the total number of trials classified as a VTE, i.e.:

$$precision = \frac{TP}{TP + FP}$$

meaning precision takes a value of 1 if all of the trials classified as a VTE are in fact VTEs, even if it does not identify all VTEs in the dataset. As a complimentary metric, recall takes FN into account:

$$recall = \frac{TP}{TP + FN}$$

and is thus a measure of how many VTEs were correctly classified out of the total number of VTEs in the dataset. For a binary classifier with equal numbers of each class, chance performance for each metric would be 0.5 on average.

We summarize performance by calculating the area under the ROC curve for each classifier. The ROC curve plots the true positive rate (recall) and false negative rate for different probability thresholds, above which the sample is assigned to the positive class (*i.e.* classified as a VTE). The false negative rate is defined as:

$$FNR = \frac{FN}{TP + FN}$$

Cross validation

To ensure our classifiers were generalizable and performance was not biased by a particular ordering of our dataset, we performed cross-validation on randomly sampled test/train splits of the dataset. For each evaluation, we used 67% of data for supervised training, and used the remainder for testing performance. For reproducibility, and to make comparisons across classifier models and feature modalities, we created a (seeded) matrix of randomly shuffled trials where each column contained a distinct ordering of trial values to use for one split of model training and testing (Liu, X.-Y. et al., 2009). For a given assessment, we used 100 distinct splits of testing and training data, giving a matrix with 100 columns. Since VTEs occur on roughly 20% of trials, every VTE in the dataset was present in each column, and a randomly drawn, equal number of non-VTEs made up the rest of the column, meaning each distinct split used the same VTE trials, but was allowed to contain different non-VTE trials (Liu, X.-Y. et al., 2009). This same matrix was used any time we evaluated classifier performance, meaning all evaluations were done using the exact same 100 iterations of test/train splits. Put another way, we assessed performance with 100 iterations of randomly selected trials constituting each test/train split, but ensured that assessments for different classifier models and feature modalities were performed on the exact same data. The figures in this paper were generated using `seed = 1`.

2.4.7 Dataset curation

Training and assessing performance of the supervised classifier required manual VTE scoring to assign labels to trials. Because it is difficult to define an exact set of criteria for scoring a VTE (hence the need for a classifier), we instead chose to have four trained raters score each trial, and used their consensus to determine the label. All raters were told to score a trajectory as a VTE if there was an indication that the rat looked toward the reward arm it did not end up choosing at least once during its trajectory. Trials where two raters scored the trajectory a VTE and two scored the trajectory a non-VTE were excluded from analysis. As shown in Supplementary Figure 2, all sessions in this dataset have an average inter-rater percent agreement above 90% and average pairwise Cohen's kappa scores above 0.7 (Hallgren, 2012; Gisev et al., 2013).

We also excluded trials based on several criteria of the hippocampal oscillation. First, we checked that the overall central tendency of the AI distribution was positive for a given session. Note that other studies have reported generally negative AIs (Amemiya and Redish, 2018; Schmidt et al., 2019). We suspect this is due to systematic shifts in theta shape characteristics across the different hippocampal axes (Buzsáki et al., 1985, 1986; Buzsáki, 2002). We also excluded trials where a 4 SD noise threshold, calculated based on the SD of the entire timeseries, was exceeded. If any session had more than 20% of its trials excluded, we did not use any of the data from that session.

2.4.8 Statistics

We performed two-sample, two-tailed Kolmogorov-Smirnov (K-S) tests to evaluate whether empirical distributions are likely drawn from the same underlying population distribution. To test whether a distribution of differences is centered at zero (i.e. to test for differences between paired groups), we performed one-sample, two-tailed Wilcoxon signed-rank tests. To assess which features exhibit statistically distinct empirical distributions when testing a number of features, we follow K-S testing with Benjamini-Hochberg (BH) false discovery rate correction to adjust p-values. Criteria for significance is set at $p = 0.01$ (1 divided by the number of iterations) for for comparing distributions and 0.05 for corrections. We also used Cohen's d metric to assess effect size, and nothe the suggestions that a value of 0.2 is considered a small effect, a value of 0.5 is considered a medium effect, and values above 0.8 are considered large effects (Sullivan and Feinn, 2012; Calin-Jageman, 2018). Effect sizes are denoted in-text by d.

2.5 References

Amemiya, S., and Redish, A. D. (2016). Manipulating decisiveness in decision making: effects of clonidine on hippocampal search strategies. J. Neurosci. 36, 814–827. doi: 10.1523/JNEUROSCI.2595-15.2016

Amemiya, S., and Redish, A. D. (2018). Hippocampal theta-gamma coupling reflects state-dependent information processing in decision making. Cell Rep. 22, 3328–3338. doi: 10.1016/j.celrep.2018.02.091

Baker, P. M., Rao, Y., Rivera, Z. M. G., Garcia, E. M., and Mizumori, S. J. Y. (2019). Selective functional interaction between the lateral habenula and hippocampus during different tests of response flexibility. Front. Mol. Neurosci. 12:245. doi: 10.3389/fnmol.2019.00245

Belluscio, M. A., Mizuseki, K., Schmidt, R., Kempter, R., and Buzsáki, G. (2012). Cross-frequency phase-phase coupling between theta and gamma oscillations in the hippocampus. J. Neurosci. 32, 423–435. doi: 10.1523/JNEUROSCI.4122-11.2012

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., et al. (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. Neuron 66, 921–936. doi: 10.1016/j.neuron.2010.05.013

Bett, D., Allison, E., Murdoch, L. H., Kaefer, K., Wood, E. R., and Dudchenko, P. A. (2012). The neural substrates of deliberative decision making: contrasting effects of hippocampus lesions on performance and vicarious trial-and-error behavior in a spatial memory task and a visual discrimination task. Front. Behav. Neurosci. 6:70. doi: 10.3389/fnbeh.2012.00070

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., et al. (2016). Prospective representation of navigational goals in the human hippocampus. Science 352, 1323–1326. doi: 10.1126/science.aaf0784

Buzsáki, G. (2002). Theta oscillations in the hippocampus. Neuron 33, 325–340. doi: 10.1016/S0896-6273(02)00586-X

Buzsáki, G., Czopf, J., Kondákor, I., and Kellényi, L. (1986). Laminar distribution of hippocampal rhythmic slow activity (RSA) in the behaving rat: current-source density analysis, effects of urethane and atropine. Brain Res. 365, 125–137. doi: 10.1016/0006-8993(86)90729-8

Buzsáki, G., Rappelsberger, P., and Kellényi, L. (1985). Depth profiles of hippocampal rhythmic slow activity ('theta rhythm'). Electroencephalogr. Clin. Neurophysiol. 61, 77–88. doi: 10.1016/0013-4694(85)91075-2

Calin-Jageman, R. J. (2018). The new statistics for neuroscience majors: thinking in effect sizes. J. Undergrad. Neurosci. Educ. 16, E21–E25. doi: 10.31234/osf.io/zvm9a

Dvorak, D., Radwan, B., Sparks, F. T., Talbot, Z. N., and Fenton, A. A. (2018). Control of recollection by slow gamma dominating mid-frequency gamma in hippocampus CA1. PLoS Biol. 16:e2003354. doi: 10.1371/journal.pbio.2003354

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn. Sci. 9, 474–480. doi: 10.1016/j.tics.2005.08.011

Fries, P. (2015). Rhythms for cognition: Communication through coherence. Neuron 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Gisev, N., Bell, J. S., and Chen, T. F. (2013). Interrater agreement and interrater reliability: key concepts, approaches, and applications. Res. Soc. Administr. Pharm. 9, 330–338. doi: 10.1016/j.sapharm.2012.04.004

Goss, A., and Wischner, G. (1956). Vicarious trial and error and related behavior. Psychol. Bull. 53:20. doi: 10.1037/h0045108

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. Tutor. Quant. Methods Psychol. 8, 23–34. doi: 10.20982/tqmp.08.1.p023

Hasz, B. M., and Redish, A. D. (2020). Spatial encoding in dorsomedial prefrontal cortex and hippocampus is related during deliberation. Hippocampus 30, 1194–1208. doi: 10.1002/hipo.23250

Hu, D., and Amsel, A. (1995). A simple test of the vicarious trial-and-error hypothesis of hippocampal function. Proc. Natl. Acad. Sci. U.S.A. 92, 5506–5509. doi: 10.1073/pnas.92.12.5506

Johnson, A., and Redish, A. D. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. J. Neurosci. 27, 12176–12189. doi: 10.1523/JNEUROSCI.3761-07.2007

Jones, M. W., and Wilson, M. A. (2005). Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. PLoS Biol. 3:e402. doi: 10.1371/journal.pbio.0030402

Kidder, K. S., Miles, J. T., Baker, P. M., Hones, V. I., Gire, D. H., and Mizumori, S. J. Y. (2021). A selective role for the mPFC during choice and deliberation, but not spatial memory retention over short delays. Hippocampus 1–11. doi: 10.1002/hipo.23306

Lever, J., Krzywinski, M., and Altman, N. (2016). Classification evaluation. Nat. Methods 13, 603–604. doi: 10.1038/nmeth.3945

Liu, X.-Y., Wu, J, and Zhou, Z.-H. (2009). Exploratory undersampling for class-imbalance learning. IEEE Trans. Syst. Man Cybern. Part B 39, 539–550. doi: 10.1109/TSMCB.2008.2007853

Malley, J. D., Malley, K. G., and Pajevic, S. (2011). Statistical Learning for Biomedical Data. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511975820

Muenzinger, K. F., and Gentry, E. (1931). Tone discrimination in white rats. J. Compar. Psychol. 12, 195–206. doi: 10.1037/h0072238

Papale, A. E., Stott, J. J., Powell, N. J., Regier, P. S., and Redish, A. D. (2012). Interactions between deliberation and delay-discounting in rats. Cogn. Affect. Behav. Neurosci. 12, 513–526. doi: 10.3758/s13415-012-0097-7

Papale, A. E., Zielinski, M. C., Frank, L. M., Jadhav, S. P., and Redish, A. D. (2016). Interplay between hippocampal sharp-wave-ripple events and vicarious trial and error behaviors in decision making. Neuron 92, 975–982. doi: 10.1016/j.neuron.2016.10.028

Redish, A. D. (2016). Vicarious trial and error. Nat. Rev. Neurosci. 17, 147–159. doi: 10.1038/nrn.2015.30

Santos-Pata, D., and Verschure, P. F. (2018). Human vicarious trial and error is predictive of spatial navigation performance. Front. Behav. Neurosci. 12:237. doi: 10.3389/fnbeh.2018.0023

Schmidt, B., Duin, A. A., and Redish, A. D. (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. J. Neurophysiol. 121, 1981–2000. doi: 10.1152/jn.00793.2018

Sullivan, G. M., and Feinn, R. (2012). using effect size-or why the p value is not enough. J. Grad. Med. Educ. 4, 279–282. doi: 10.4300/JGME-D-12-00156.1

Tolman, E. C. (1938). The determiners of behavior at a choice point. Psychol. Rev. 45, 1–41. doi: 10.1037/h0062733

Voss, J. L., and Cohen, N. J. (2017). Hippocampal-cortical contributions to strategic exploration during perceptual discrimination. Hippocampus 27, 642–652. doi: 10.1002/hipo.22719

Chapter 3

A selective role for the mPFC during choice and deliberation, but not spatial memory retention over short delays

A version of this chapter has been published in the journal *Hippocampus*.

Citation:
Kidder KS, Miles JT, Baker PM, Hones VI, Gire DH, Mizumori SJY (2021). A selective role for the mPFC during choice and deliberation, but not spatial memory retention over short delays. Hippocampus. doi: 10.1002/hipo.23306

Abstract

Important interactions between memory and decision-making processes are required to maintain high-levels of spatial working memory task performance. Past research reveals that the medial prefrontal cortex (mPFC) and hippocampus (HPC) are both vital structures involved in these processes. Recent evidence suggests that interactions between these two structures are dynamic and task dependent. However, there exists uncertainty surrounding the specific conditions that recruit mPFC contributions to these tasks, specifically regarding its role in retaining information online during delay periods. To address this issue, we tested rats on a spatial-delayed alternation task in which we utilized a closed-loop optogenetic system to transiently disrupt mPFC activity during different task epochs (delay, choice, return). By analyzing the effects of mPFC disruption on choice accuracy and a deliberative behavior known as vicarious-trial-and-error (VTE), our study revealed several interesting findings regarding the role of the mPFC in spatial-working memory tasks. The main findings include: 1) choice accuracy in the spatial-delayed alternation (SDA) task was impaired when the mPFC was disrupted during the choice epoch and not delay or return epochs, 2) mPFC disruption resulted in a non-epoch specific reduction in VTE occurrence which correlated with impairments in task performance. Taken together, findings from this study suggest that, during spatial decision-making, contributions made by the mPFC are specific to points of deliberation and choice (not delay), and that VTEs are a deliberative behavior which relies on intact mPFC function.

3.1 Background and Introduction

When faced with uncertainty, animals must evaluate and manipulate past memories to bias future behaviors that increase the likelihood of obtaining desired outcomes. Decades of human and animal research have implicated the hippocampus (HPC) and medial prefrontal cortex (mPFC) as two important structures involved in the brain's capacity to make effective decisions. The HPC is known for its role in the initial storage and retrieval of declarative and episodic memory, while the mPFC is known for its role in outcome evaluation, response inhibition, implementation of task rules or strategies, and several other higher order executive functions. Together, these two structures are proposed to be part of a decision-making and working memory (WM) circuit that facilitates interactions of recent memory with task rules and strategies to aid in deliberation and choice selection.

Anatomical studies have revealed these two structures are connected directly through a ventral HPC (vHPC) to mPFC pathway, and indirectly through a bidirectional nucleus reuniens (RE) mediated pathway that links the mPFC with dorsal HPC (dHPC) (Dolleman-van der weel et al., 2019, Thierry et al., 2000). Numerous disconnection and lesion studies demonstrate the importance of HPC-mPFC communication during spatial working memory (SWM) and decision-making tasks (Xia et al., 2019, Goto & Grace, 2007, Ito et al., 2015, Maharjan et al., 2018). For example, Avigan et al. (2020) found that contralateral, but not ipsilateral, inactivation of the mPFC and either the dHPC or vHPC impaired WM performance. Another study by Floresco et al. (1997) found that contralateral disconnections of HPC and PFC impaired performance on a spatial task that had a delay, but not on the same task without a delay. Thus, there is strong evidence from disconnection and lesion studies that interactions between the mPFC and both HPC regions are vital to WM and decision-making processes; however, questions regarding the mPFC's specific functional contributions to these processes remain.

Electrophysiological recordings from the HPC and mPFC further support the notion that these structures interact during WM tasks. By recording oscillatory activity in dHPC and mPFC as animals performed a WM task, Hallock et al. (2016) revealed a significant increase in theta coherence between these two structures during the choice epoch of this task, and not at the start box or stem areas. Similarly, Jones and Wilson (2005) reported increased HPC-mPFC (4-12Hz) oscillatory coherence at the choice-point of a WM task, while also finding that mPFC single-units displayed increased spatial information during task epochs where significant phase-locking of mPFC units to HPC theta was observed. Numerous studies have shown similar results regarding single-unit and oscillatory interactions between HPC and mPFC (Benchenane et al., 2010, Colgin, 2013, Fell & Axmacher, 2011, Spellman et al., 2015, Tamura et al., 2017), suggesting these two structures interact in dynamic, task-dependent ways to facilitate high-level WM and decision-making performance. These findings depicting dynamic communication between HPC and mPFC are not surprising given that decision-making is defined as the process of selecting an action based on important interactions between memory and decision systems (Redish & Mizumori, 2015).

While it's clear that the HPC and mPFC interact during memory-guided decision-making tasks, it's unclear how mPFC activity relates to the dynamics of the task. For example, several studies (Baeg et al., 2003, Fuster & Alexander, 1971, Zylberg & Strowbridge, 2017) have described mPFC units with elevated or sustained firing activity during delay periods of WM tasks which correlated with WM performance. Kamigaki & Dan (2017) also found that mPFC disruption during the delay period of a go-no-go task significantly impaired task performance. These findings led many to suggest the mPFC plays a role in maintaining important task relevant information online during delay periods (Preston & Eichenaum, 2013). Seemingly contradictory to these findings are reports (Hyman et al., 2010, Pratt & Mizumori, 2001) of sparse, if any, delay-related spatial activity by mPFC units, suggesting a more nuanced role in WM performance. Also, while recording mPFC single-unit activity during a spatial delayed alternation task, Horst and Laubach (2012) were unable to find mPFC neurons that persistently fired in a spatially selective manner throughout the delay. These authors did find, however, that many mPFC cells were responsive to the outcome of upcoming decisions, supporting an alternative view to the common notion that the mPFC is involved in the online storage of information. Rather, the mPFC may contribute to the prospective organization of information to guide future behavior, a function also attributed to the HPC (Guise & Shapiro, 2017).

Deliberation of path options ultimately aids the selection of the best prospective sequence that will lead to a desired outcome. It is proposed that these deliberative cognitive processes may be reflected in a behavior known as *vicarious trial and error* (VTE) (Muenzinger, 1938, Redish, 2016). VTEs involve a back-and-forth sweeping of an animal's head and/or body between possible options before finally making a choice. This behavior most often occurs when animals are faced with uncertainty, such as when presented with multiple options on a maze, and it seems that VTE occurrence increases with decision difficulty (Schmidt et al., 2013, Schmidt et al., 2019). This evidence suggests that VTEs are a behavior which involve HPC-mPFC interaction and additionally, a study by Santos-Pata et al. (2018) found that humans seem to use similar head scanning behaviors for deliberation. Therefore, studying VTEs has the potential to reveal insights into the functional contributions of these two structures and their interactions in decision-making and WM processes.

There exists a wealth of strong correlative evidence supporting the notion of task-dependent HPC-mPFC communication. However, past studies have lacked the ability to manipulate these structures on time scales that are relevant to the temporal dynamics of cognitive processes, and as such have limited our ability to causally test mPFC task-dependent contributions to decision-making and WM processes. To address this issue, we used a spatial delayed alternation (SDA) task and a closed-loop optogenetic system to transiently disrupt mPFC activity in rats as they traversed different epochs of the task (delay, choice, return). We hypothesized that if the mPFC was preferentially involved in retaining recent spatial memory online, then mPFC disruption during the delay epoch should impair SDA performance. On the other hand, if the mPFC was preferentially involved in choice and/or deliberation, then mPFC disruption during the choice epoch should impair performance and reduce the occurrence of VTEs. We did not expect performance impairments after return epoch disruption due to the fact that our well-learned task does not require flexible rule switching between trials. Results from this study importantly discern a selective role for the mPFC during spatial delay responding by showing that it is necessary for deliberation and choice behaviors, but not for retaining information online over short delays.

3.2 Methods

3.2.1 Animals

12 Long-Evans rats (male = 3, female = 9, Charles River Laboratories) were used in this experiment, 3 of which were used as control animals. Animals were housed in a temperature-controlled environment with a 12 h light/dark cycle, and all experiments were conducted during the light phase. All animals were given food and water ad libitum and handled for at least 5 days before maze training began. During training and testing, rats were maintained at ~85% of their maximum free feeding body weight. All animal care was conducted according to guidelines established by the National Institutes of Health and approved by the University of Washington's Institute for Animal Care and Use Committee (IACUC).

3.2.2 Apparatus

Testing took place in a sound attenuated room adjacent to an external room that contained all electronic and recording devices. Inside the testing room, the maze was encircled by black curtains that extended from the ceiling to the floor and had various visually-distinct shapes attached to it that could be used as landmarks for the rat. The maze was a black plexiglass elevated cross-maze (arms 58 X 5.5cm, elevated 80cm from the floor). In order to control rats' behavior, maze arms were controlled via the use of arduinos to LabVIEW 2016 software (National Instruments, Austin, TX, USA). Two arms designated east (E) and west (W) contained 3D printed food wells connected to computer-controlled pellet dispensers (Lafayette Instruments, Lafayette, IN, USA). The north (N) and south (S) arms were used as start arms in the task. Each maze arm was hinged midway so that the proximal end could be raised and lowered via servos connected to arduino boards.

3.2.3 Surgeries and electrophysiological recording procedures

Shortly after arriving at our facility rats were anesthetized using 1.0%–2.0% isoflurane in oxygen (flow rate 1.0 L/min) and placed into a stereotaxic apparatus (KOPF). Rats then underwent surgery involving bilateral mPFC (AP: 3.0mm, ML: ± 0.8, DV: -3.5) intracranial injections of the excitatory optogenetic viral construct AAV5-CaMKlla-hChR2-mCherry (Addgene: CS1096). Following surgery, rats were allowed approximately seven days of recovery before beginning maze training procedures.

Upon reaching performance criterion of three consecutive days of at least 80% correct choices on the spatial-delayed alternation task, rats underwent micro-drive implantation surgery. Each 3D printed (Formlabs) micro-drive consisted of both recording tetrodes and optic fibers. 14-16 gold plated tetrodes (nichrome, SANDVIK) were implanted unilaterally (hemispheres counterbalanced between animals) into the CA1 region of dHPC (AP: -3.0, M/L: ±2.0mm, D/V: -1.8mm). Two optic fibers (one per hemisphere) were implanted bilaterally into the mPFC. Tetrodes were connected to a 64-channel EIB containing two, 32-channel Omnetics connectors that were connected to OpenEphys acquisition boards (open-ephys.org). To eliminate external noise, drive enclosures were created with a 22mm plastic tube that was lined with nickel coated aluminum foil. One ground wire connected the foil lined tube with the EIB and, during surgery, another ground wire was implanted near the cerebellum just below the skull. After surgery, rats were allowed to recover for approximately seven days before entering the next phase of testing. As testing resumed, HPC tetrodes were lowered between 20-80 microns a day until LFP signatures revealed proper placement into the pyramidal layer of HPC-CA1.

3.2.4 Spatial delayed alternation (SDA), training, and experimental design

Training procedures were similar to those previously published (Baker et al. 2019). Briefly, prior to surgery, rats were food restricted to 85% of their free feeding weight. Once their weights were stable, rats were trained to alternate between two oppositely positioned reward arms in a plus maze in order to receive reinforcement of two 45 mg sucrose pellets (TestDiet, Richmond, IN, USA). These reward arms were designated E and W arms while the other arms (N and S) were used as start arms for the task.

Training

Initial training was performed to acclimate rats to the general organization of a trial. These consisted of 10 forced choices alternating between choice arms, followed by 35 free choice trials in which both arms were available to choose. A given trial consisted of the rat waiting in a start arm (N or S) for five seconds. Once this inter-trial interval was completed, either one (in forced trials) or both (in choice trials) goal arms were raised. Any choice in this phase led to the delivery of a reward. Once a choice was made, the reward was delivered, the opposite choice arm and the start arm were lowered, and after 2.5 sec a start arm was raised to allow the rat to return to start the next trial. The start arms were pseudo-randomly chosen with no more than two consecutive starts from the same arm. Once rats were able to complete 45 trials in 30 min or less for three consecutive days, they were then trained on the SDA task.

In the SDA task (**Figure 3.2B**), the inter-trial interval (delay) was increased to 10 sec and rats were only rewarded if they selected the choice arm opposite from the previous trial. If the same choice was repeated (e.g. W and W) then no reward was given. Additionally, no forced choice trials were offered during the SDA task. All other aspects of a trial were the same as in the prior training phase. Once rats were able to complete 60 trials with at least 80% of choices being correct (i.e. 80% choice accuracy) for three consecutive days, rats were placed on free feed in preparation for micro-drive surgery. Following recovery from surgery, rats were again run on the SDA task to ensure retention of the task prior to moving to the experimental phase. If rats were able to perform the task on two consecutive days with at least 80% choice accuracy, they advanced to the experimental phase of the task. *See supplemental figure S-1 for SDA training data.*

Experimental design

In order to examine contributions made by the mPFC at different points throughout decision-making and WM processing, we split the SDA task into 3 distinct epochs: delay (to test working memory function), choice (to test decision-making), return (to test reward acquisition; figure 2B). During the delay epoch animals were isolated on either start arm for 10 sec. At the end of the delay, maze arms were raised and the choice epoch began. The choice epoch in our task was defined as the 5 sec directly following the end of the delay. The return epoch began from the time of reward consumption and lasted for 10 sec. Animals were not allowed to travel to the next trial's start arm until the 10 sec following reward consumption had elapsed.

After rats achieved post-surgical asymptotic performance on the SDA task, they underwent a counterbalanced series of 3 testing conditions that occurred on different days. Optogenetic stimulation was applied to the mPFC during only one epoch per condition. Each test session consisted of 60 trials which were split into two blocks of 30 trials each (baseline and stimulation). Rats had up to two repeats of each stimulation condition, for a maximum of 6 testing days per animal. Each day was considered an individual data point in our samples. There was a small subset of trials where, while performing choice epoch stimulation conditions, rats waited in the delay arm until stimulation had stopped. Any trials where this happened were removed. If an animal waited on more than 50% of the trials during the stimulation block, that session was removed from analysis. Five rats were exclusively tested with HPC-theta driven stimulation, 2 were exclusively tested with 20Hz stimulation, and 1 rat was tested with both types of stimulation (**Figure 2.2A**)(see *Optogenetic stimulation* for more information). Animals tested in 20Hz conditions underwent one round of testing conditions. Overall, our experiments resulted in 40 sessions which met inclusion criteria, and these were used in subsequent analyses.

3.2.5 Data Acquisition

Behavior Tracking

Rat locations were determined by subtracting the previous frame from a background average taken at the beginning of each session. Pixels that showed an above threshold difference in brightness were identified and used to track movement of the rat based on proximity to the previously identified location. Position analysis was performed using a custom LabView (National Instruments, Austin, TX, USA) routine detecting LED's attached to a tether coming out of the rat's micro-drive. Camera frames were recorded at approximately 35 Hz using a tracking camera (SONY). Frames were time-stamped with a millisecond timer run by LabView and sent to the OpenEphys acquisition software (open-ephys.org) for later alignment of electrophysiological and position information.

Electrophysiology

Electrophysiological data were sampled at 30 kHz using Intan RHD2164 headstages connected to an OpenEphys 64-channel electrode interface circuit board, and acquired with an OpenEphys acquisition board (Intan, RHD2000), all of which were available through open-ephys.org. For closed-loop HPC theta-based stimulation of the mPFC, incoming signals were band-pass filtered at 4-12 Hz using the OpenEphys GUI's built-in Bandpass Filter module. The peak or trough of the ongoing HPC theta rhythm was detected using the Phase Detector module, set to either the ascending or descending phases, respectively, of HPC theta. Ascending or descending phases were chosen in order to accommodate the slight processing delay of approximately 15ms in the system.

Optogenetic stimulation

To stimulate opsins and disrupt the mPFC, we used a 473nm laser (Laserglow technologies) held to a power of approximately 7mW for each animal, measured at the tip of the optic fiber just before implantation. Fiber-optic cables attached to the power source and were affixed to fibers targeting the mPFC of rats before each testing session. Two stimulation patterns were used for our experiments. For rats undergoing open-loop stimulation, laser pulses were triggered by rat position and occurred at 20 Hz, with 50% duty cycle, for the duration of a given behavioral epoch (delay, choice, or return). Rats undergoing closed-loop experimentation received stimulation triggered by epoch and gated by ongoing HPC theta phase (see *Electrophysiology*). Phase-based laser signals were sent to the laser's power source by an Arduino (UNO-R3) that interfaced with OpenEphys software. A second Arduino interfaced with LabView software to gate the laser activation by maze position (epoch). *See figure S-2 for examples of mPFC cellular responses to ChR2 stimulation.*

Vicarious Trial and Error

Vicarious trial and error (VTE) behavior occurs when rats reach a decision point and deliberate possible options by sweeping their head and/or body back-and-forth between options before ultimately making a decision (Redish, 2016). VTEs were manually and independently scored for each trial of each experimental day by five trained Mizumori lab members. At least 4 raters had to agree on the outcome of a trial (VTE or non-VTE) in order for the trial to be included in further analyses. *For example traces of VTEs, see supplemental figure S-3.*

3.2.6 Histology

After the completion of all testing sessions, tetrode and optic fiber locations were verified with marking lesions. Rats were deeply anesthetized with 4% isoflurane, and each tetrode tip location was marked by passing 9 μA current through each tetrode wire for 10 seconds. Animals were then given an overdose of sodium pentobarbital and transcardially perfused with 0.9% saline and a 10% formaldehyde solution. Brains were stored at 4°C in 10% formalin for 1 day followed by 4 days in a 30% sucrose solution. The brains were then frozen and cut into coronal sections (40 μm) on a freezing microtome. HPC sections were mounted on gelatin-coated slides, stained with cresyl violet, and examined under a light microscope. mPFC sections were mounted onto slides and then imaged using a fluorescent microscope to confirm viral expression.

3.3 Results

3.3.1 Histology

Tips of HPC tetrodes and mPFC optic fibers were located in the targeted brain areas (**Figure 3.1**). Most animals had HPC tetrodes terminating in the stratum oriens and pyramidal layer of CA1. A minority of tetrode bundles terminated near the CA1 fissure. Optic fibers were bilaterally located in the mPFC. Optic fiber tip placement was evenly distributed along the D/V axis of the prelimbic cortex starting from the ACC-PL border to the PL-IL border. Viral expression in the mPFC surrounded all optic fiber tips. Viral expression was relatively equal between mPFC hemispheres except for two animals with unilateral expression. The latter animals' data were then considered as a control (see below).

3.3.2 mPFC disruption impaired SDA-task performance regardless of stimulation type

In order to test mPFC involvement in the SDA task, and to determine if the different types of optogenetic stimulation produced different behavioral effects, we used a two-way within-subjects ANOVA to compare choice accuracy scores between baseline and stimulation blocks for 20Hz versus theta-based stimulation types (figure 2.3A).

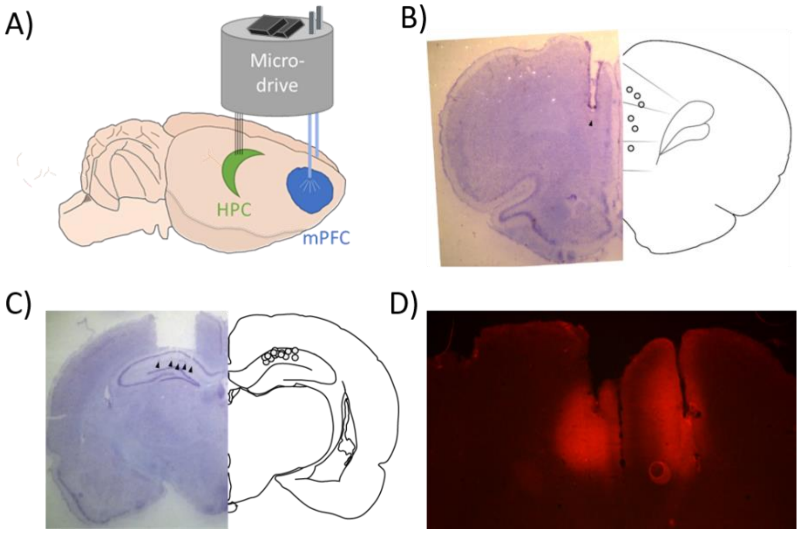

Figure 3.1: (A) Illustration of optogenetic micro-drives used in this study. Approximately 14 tetrodes were implanted unilaterally into CA1 of HPC. One optic fiber per hemisphere was extended through the drive into the PFC. **(B)** (left) Example mPFC section showing optic fiber placement in mPFC and (right) optic fiber placements of fiber tips from all rats. **(C)** (left) Example HPC section showing tetrodes terminating near the CA1 fissure. Black arrows represent tetrode tips. (Right) HPC tetrode placements in all rats. **(D)** Example of mPFC Channelrhodopsin-2 viral expression surrounding optic fiber tips, detected using fluorescent imaging of m-Cherry.

Using a two-way within-subjects ANOVA, the mean choice accuracy scores for baseline ($x = 95.1\%$, $SD = 5\%$) and stimulation ($x = 84.8\%$, $SD = 11.6\%$) blocks were significantly different from each other ($F(1, 76) = 25.92$, $p < 0.05$), the mean choice accuracy scores between 20Hz ($x = 91.5\%$, $SD = 9.3\%$) and theta-based ($x = 89.5\%$, $SD = 10.6\%$) stimulation types were not significantly different from each other ($F(1, 76) = 0.70$, $p > 0.05$), and there was no significant interaction between block type and stimulation type ($F(1, 76) = 0.18$, $p > 0.05$). Our study was not sufficiently powered to look for sex differences in SDA performance, however, we analyzed the means and standard deviations between the two sexes and saw that for all scores there were no apparent sex differences in performance or in deficits caused by stimulation.

To test for the possibility that laser stimulation impaired task-performance by inadvertently distracting animals, we analyzed data from 5 choice epoch sessions which were obtained from 3 control animals (2 animals with unilateral ChR2 expression and 1 animal in which optic fiber tips missed the target area). A one-way within-subjects ANOVA revealed no differences in choice accuracy between baseline ($x = 90.7\%$, $SD = 5.3\%$) and stimulation blocks ($x = 90.0\%$, $SD = 6.3\%$) for these animals ($F(1,8) = 0.03$, $p > 0.05$). Additionally, using a paired samples t-test, we found that the time to completion (defined as the time it took an animal to travel from the start arm to the reward arm of their choice) was not significantly different between baseline and stimulation blocks ($t(39) = 4.90$, $p > 0.05$), suggesting that optogenetic mPFC stimulation did not increase or decrease animals' motivation during the SDA task.

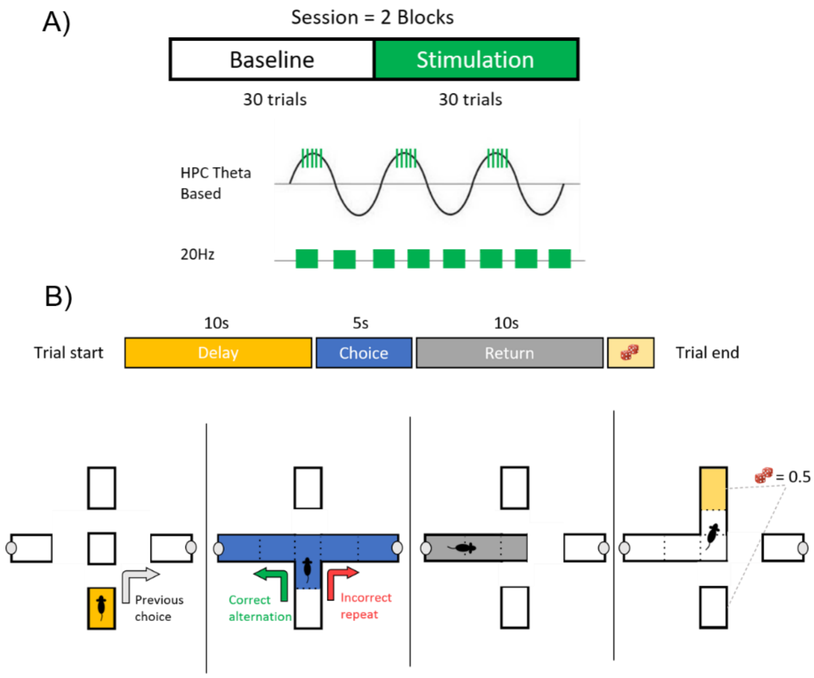

Figure 3.2: (A) (top) Each session consisted of 60 trials that were split into two blocks (Baseline and Stimulation) of 30 trials each. (bottom) Illustrations of the two ChR2 light stimulation parameters used in this experiment. HPC Theta based stimulation involved the online detection of specific phases (peak/trough) of the HPC theta cycle. Upon theta phase detection and detection of the conditions particular stimulation epoch, 5 light pulses at 100HZ (50% duty cycle) were sent to the mPFC to disrupt it. The other stimulation type involved constant 20Hz (50% duty cycle) light pulses during the duration of a conditions particular stimulation epoch. For each session, one of the three epochs (see below) was selected in which laser stimulation was gated to only occur during that epoch for all of that sessions stimulation block trials. **(B)** (top) Schematic showing task progression within a single trial. The delay epoch lasted for 10 sec, the choice for 5 sec, and the return lasted for 10 sec. At the end of the 10 sec return epoch the next trial's start arm was pseudo-randomly selected (dice) at which point the chosen maze arm was raised to allow the rat to enter that start arm, which then initiates the next trial's delay epoch. (bottom) Schematic of the spatial delayed alternation (SDA) task which takes place on an automated plus-maze. Start arms are opposite from each other to the north and south, while the reward arms are opposite from each other to the east and west.

These results suggest that mPFC disruption impaired SDA task performance regardless of stimulation type. Since there was an identical pattern of choice accuracy effects between 20Hz and theta-based simulation conditions, we combined data across animals with the different stimulation types in subsequent analyses.

3.3.3 SDA performance impairments result from mPFC disruption during the choice-epoch

To determine if mPFC disruption selectively impaired performance in any of the 3 epoch conditions (delay, choice, return) of the SDA task, for each session we compared the change in choice accuracy by subtracting each session's stimulation block choice accuracy from its baseline block's choice accuracy (**Figure 2.3B**). A one-way ANOVA revealed a significant effect of epoch on change in accuracy scores ($F(2, 37) = 25.17$, $p < 0.05$), indicating that mPFC disruption had differential effects on choice accuracy depending on the particular stimulation epoch. Post hoc comparisons using the Tukey HSD test indicated that the mean change in choice accuracy for the choice epoch condition ($x = -21.1\%$, $SD = 8.6\%$) was significantly greater than that observed for delay (x = -6.7%, SD= 7.1%) and reward epoch conditions (x = -2.4% , SD = 5.5%), the latter of which were not significantly different from each other. Results from this test indicate that mPFC disruption significantly reduced choice accuracy when mPFC disruption occurred while rats traversed the choice epoch, and not when the mPFC was disrupted during delay or return epochs.

3.3.4 mPFC disruption resulted in a non-epoch specific reduction in VTE occurrence

We next analyzed the effect of mPFC disruption on the proportion of VTEs that occurred. Animals on average performed VTEs during 26.9% (SD = 10.7%) of trials during baseline blocks and 18.7% (SD = 13.4%) of trials during stimulation blocks. The means of the two blocks were significantly

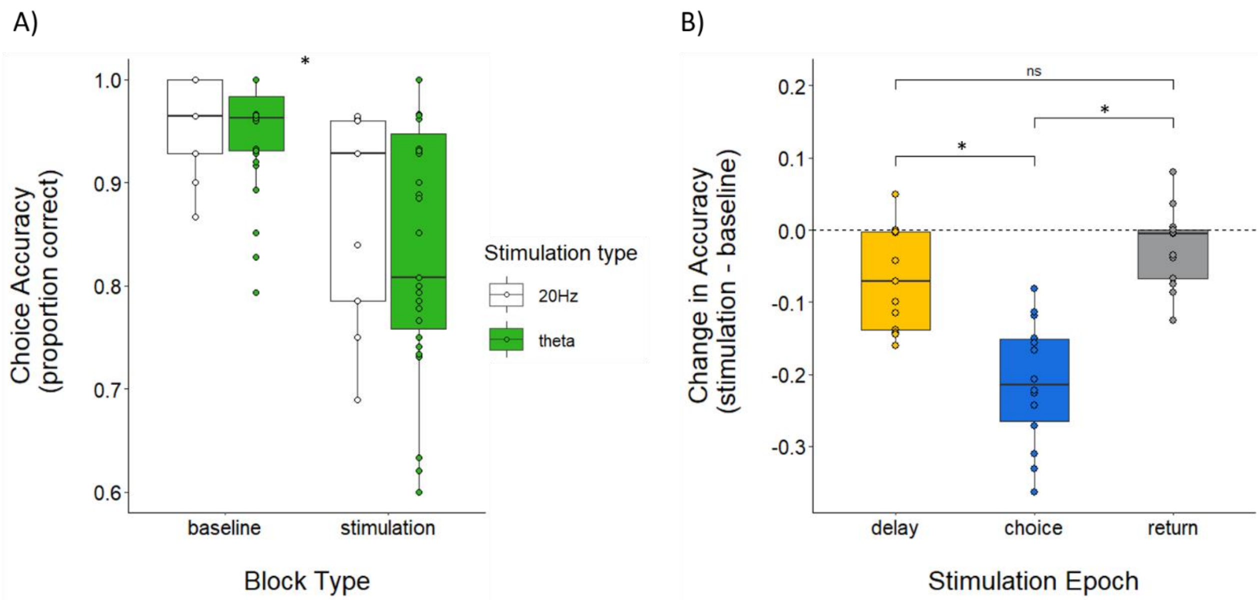

Figure 3.3: (A) Boxplots showing choice accuracy scores during the SDA task between baseline and mPFC stimulation blocks. Sessions are split by stimulation type, with 20Hz in white and HPC theta based in green. **(B)** Boxplots showing changes in accuracy following mPFC disruption in each of the 3 epoch conditions. Sessions with different stimulation types were combined in subsequent analyses. Change in accuracy scores were calculated by subtracting choice accuracy in the stimulation block from choice accuracy in the corresponding session's baseline block. The dashed line represents no change in choice accuracy between blocks. (*p< .05)

different from each other (t(39) = -3.288 , p < 0.05), with an average decrease in VTE occurrence of 8.2% between baseline and stimulation blocks (**Figure 3.4A**).

Furthermore, to determine whether the change in VTE occurrence due to mPFC disruption was selective to any of the 3 epochs, we calculated the change in VTE occurrence by subtracting the VTE occurrence in each sessions stimulation block from the VTE occurrence in its baseline block (**Figure 3.4B**). Using a one-way ANOVA we compared the change in VTE occurrence between each of the 3 epoch conditions. The initial analysis revealed a significant effect of epoch on the change in VTE occurrence scores (F(2, 37) = 3.34 , p < 0.05) suggesting that there were fewer VTEs when the mPFC was disrupted in choice epoch conditions. However, this effect did not survive *post hoc p*-value corrections for multiple comparisons. The Tukey HSD test revealed that the mean changes in VTE occurrence for the choice-epoch condition (x = -16.6%, SD = 16.9%), delay condition (x = -3.1%, SD= 14.5%) and reward condition (x = -4.4% , SD = 13.4%) were not significantly different from each other. Therefore, results from these tests show a generalized decrease in VTE occurrence due to mPFC disruption that was not specific to any of the 3 epochs, although there was a trend towards reduced VTEs in the choice epoch.

3.3.5 Decreased VTE occurrence following mPFC disruption correlates with impaired SDA performance

To determine if the occurrence of VTEs was related to choice accuracy, a Pearson product-moment correlation coefficient was computed to assess the relationship between the change in VTE occurrence and the change in choice accuracy that resulted from mPFC disruption (figure 5). We found a significant positive correlation between these two variables ($r = 0.416$, $n = 40$, $p < 0.05$) suggesting that as animals' choice accuracy decreased in the SDA task (after mPFC disruption) they had a tendency to perform fewer VTEs.

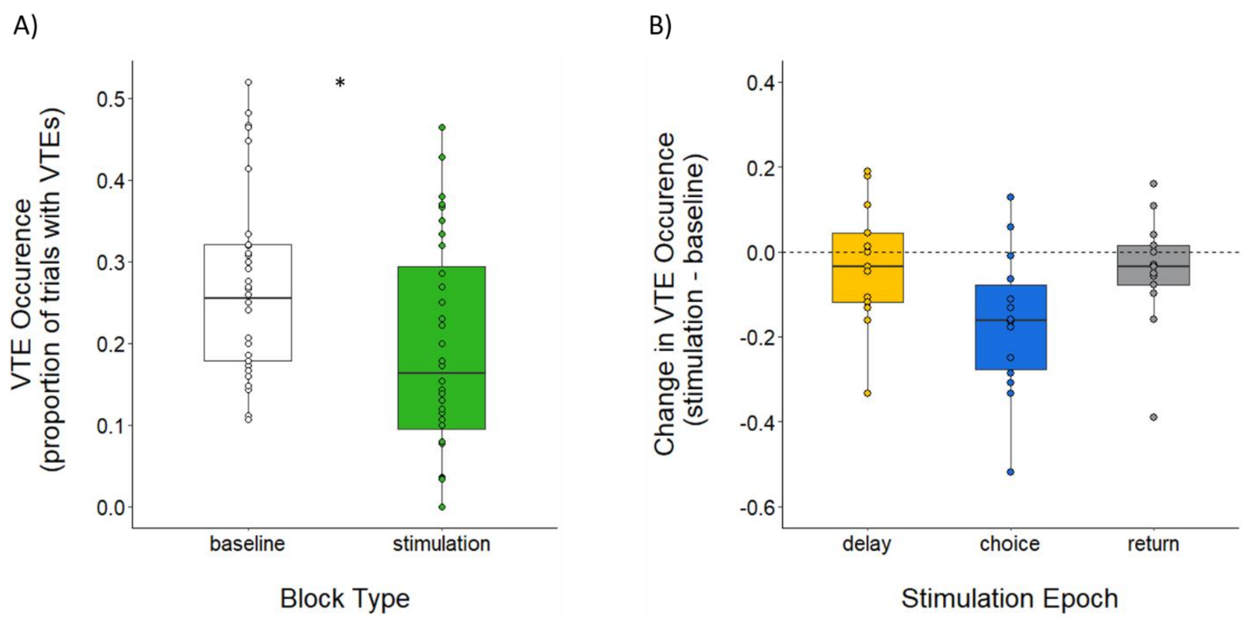

Figure 3.4: (A) Boxplots showing the proportion of VTEs occurring in each block type. (*p< .05) **(B)** Boxplots showing the change in VTE occurrence in each epoch condition. The change in VTE occurrence was calculated by subtracting the stimulation block's VTE occurrence from the baseline block's VTE occurrence. Dashed line represents no change in VTE occurrence between blocks.

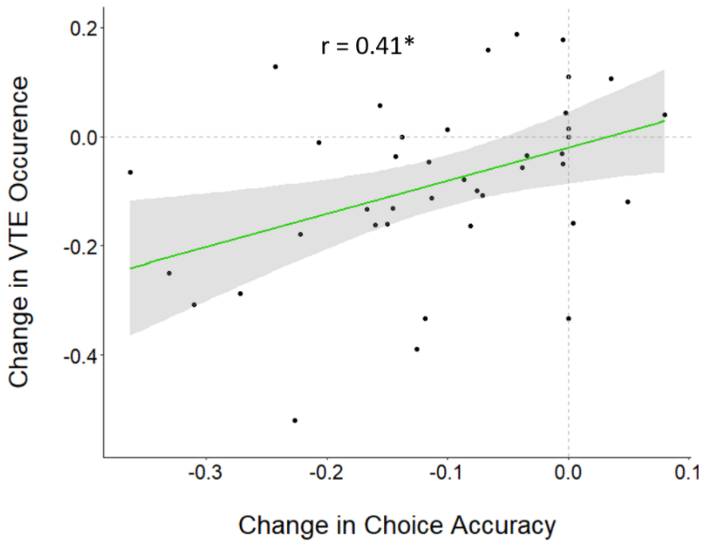

Figure 3.5: Analysis revealed a significant positive correlation (r=0.41, * p<.05) between the change in choice accuracy and change in VTE occurrence. Shaded area represents 95% CI and dotted grey lines represent points where there is no difference in scores between blocks.

3.4 Discussion

Consistent with past studies, we found that bilateral mPFC disruption impaired SWM performance in the SDA task. In addition, our results revealed that mPFC disruption impaired choice accuracy only when the mPFC was disrupted at the choice epoch, and not when it was disrupted during delay and return epochs. This effect held regardless of whether the stimulation was a constant 20 Hz stimulation, or stimulation patterned after the ongoing dHPC theta. We also found that although mPFC disruption reduced the occurrence of VTEs during the SDA task, regardless of when the mPFC was disrupted, there was a strong trend towards fewer VTEs occurring after mPFC disruption in the choice epoch. Finally, we observed that reduced VTE occurrences correlated with decreased choice accuracy following mPFC disruptions. Control experiments demonstrated that the behavioral impairments were not a result of unwanted effects of laser stimulation, such as serving as a visual distraction. Taken together, our findings show an epoch specific role for the mPFC in SWM tasks, suggesting a select role of the mPFC in initiating deliberative processes related to choice outcomes.

3.4.1 SDA performance impairments dissociate mPFC's role in choice behavior versus spatial working memory retention over a short delay

The mPFC and HPC both have known roles in SWM and decision-making, and there is strong correlative evidence which suggests HPC-mPFC communication is critical during tasks involving these processes (Guise & Shapiro, 2017, Zielinski et al., 2017). For example, studies reveal increased theta coherence between these structures when animals make decisions during SWM tasks (Benchenane et al., 2010, O'Neill et al., 2013, Hallock et al., 2016), as well as increased cross-correlated firing of HPC-mPFC neuronal pairs, and increased entrainment of mPFC single-unit activity to HPC theta (Hyman et al., 2010, Jones & Wilson, 2005, Paz et al., 2008). Based on these reports, we proposed that mPFC disruption at least during the choice epoch, and possibly at the delay epoch (see below), should decrease choice accuracy in our SDA task. While we did not analyze HPC-mPFC neural communication in the present study, electrophysiological evidence provided from past studies suggests this communication should have increased in our task as animals traversed the choice epoch. Therefore, the observed impairments in choice accuracy when the mPFC was disrupted during the choice epoch could likely be due to a disruption in the communication between HPC and mPFC.

Another major line of research suggests that HPC-mPFC communication may be important for maintaining information online during delay periods; several studies have found that subsets of PFC neurons show delay related activity that correlates with WM performance (Batuev et al., 1979, Warden & Miller, 2010, Zylberg & Strowbridge, 2017). Seemingly at odds with this evidence, the present study found no significant impairments in SWM performance when the mPFC was disrupted during the delay epoch. One explanation for our finding is that redundant information could be shared between HPC and mPFC (Lee & Kesner, 2003). Results from a study by Churchwell and Kesner (2011) suggest that over a short delay of 10sec, but not over a long delay of over 5min, spatial information carried by either structure may be sufficient to successfully complete SWM tasks. This means that while the mPFC could maintain spatial information about past trials online, disrupting mPFC information during a short delay is not sufficient to have an effect on performance since the HPC's memory functions remain intact to solve the task.

Units in the mPFC appear to carry a diversity of non-spatial information such as choice outcome and reward magnitude (Baeg et al., 2003, Pratt & Mizumori, 2001, Zylberg & Strowbridge, 2017). Such information is thought to importantly inform flexible response selection functions of the mPFC. In our study, disrupting the mPFC during the return epoch did not result in impaired SWM performance, indicating that information about choice outcome or rewards carried by the mPFC is not necessary when performing a SWM task. Since non-spatial information was not required to solve the SDA task it is possible that the mPFC is needed during the return epoch in other tasks that require the use of such non-spatial information and it is also possible this information could be retained online by the mPFC during the delay epoch. This leaves an open avenue for future research that tests the role of the mPFC in retaining and utilizing its non-spatial information throughout different phases of WM tasks.

A major unresolved question is the identity of the specific types of information that were perturbed when the mPFC was disrupted during the choice epoch. In our task, the choice epoch is functionally complex: it includes the time when animals sample spatial cues as they encroach upon the choice point, recall of the previous reward location, deliberation of choice options based on prior experience, choice selection, and the sampling or encoding of locations shortly after the choice was made. Disruption of any one or a combination of these functions could have led to choice inaccuracies. Spellman et al. (2015) demonstrated similar functionally complex interpretations of performance changes in an analogous delayed non-match to place task. Inhibition of vHPC-mPFC terminals decreased accuracy only when inhibition occurred during the sample phase (trial N) and not during the choice phase (trial N + 1). This led the authors to hypothesize that performance deficits came from impairing vHPC's ability to update the mPFC with the animal's spatial location during the sample phase. The choice epoch in our SDA task combines the sample and choice phases from Spellman et al.'s study, so the interpretation that our performance deficits could have resulted from a similar failure of the mPFC to properly encode animals' past location is plausible. When taking our VTE findings into account, however, a more general interpretation may be that our stimulation disrupted the mPFC's ability to engage the HPC in memory-guided deliberation (Schmidt et al., 2019, Redish, 2016). Theories surrounding VTEs suggest that this behavior scales with the level of uncertainty as a compensatory mechanism to aid in deliberation (Schmidt et al., 2019). If mPFC disruption prevented encoding of the animals past location, we would expect uncertainty in the task to increase, which should be accompanied by an increase in VTEs. In contrast, we saw decreases in the number of VTEs upon mPFC stimulation. This suggests that in the present study, reductions in choice accuracy may have been due to an inability of the mPFC to engage deliberative behaviors rather than disrupting the encoding of past locations.

Further evidence of behaviorally and cognitively complex contributions by the mPFC was provided by Guise and Shapiro (2017) and others (Dolleman-van der weel et al., 2019, Goto & Grace, 2017). These studies suggest that the mPFC sends signals relating to task rules or strategies to the HPC via a bi-directional pathway involving the nucleus reuniens. These signals are thought to become incorporated into HPC representations which support deliberation. Such an interpretation of the functional significance of mPFC output to the HPC suggests that in our task, disruption of important mPFC task rule information normally signaled to the HPC during the choice epoch may have resulted in the observed impairments to choice accuracy.

Future experiments must rely on innovative experimental designs to tease apart the multitude of cognitive operations that support deliberation and choice behavior. Nevertheless, when considered together, our findings of SWM deficits only when the mPFC was disrupted during the choice epoch suggest that the mPFC is involved in deliberative processes that occur during critical decision-making time points, and not in the process of retaining spatial information online during a short delay.

3.4.2 VTEs reflect deliberation and are influenced by processes that require intact mPFC function

VTEs are theorized to be the behavioral correlate of deliberation. As such, current models of VTEs suggest this behavior manifests as the mPFC helps bias or generate prospective path options in conjunction with HPC. In this way, the back-and-forth sweeping of the animal's head between possible routes is thought to be reflective of the animal generating and evaluating possible options and outcomes (Amemiya & Redish, 2016, Redish, 2016). Furthermore, it is believed that this behavior could sometimes act as a compensatory mechanism which occurs when information regarding past memories is uncertain (Papale et al., 2010).

In support of this model, recent work from the Redish lab (Schmidt et al., 2019) revealed several interesting findings: 1) In their restaurant row foraging task they observed animals performing more VTEs as decision difficulty increased, 2) mPFC inactivation using DREADDs reduced the occurrence of VTEs, and 3) mPFC inactivation impaired the number of HPC theta sequences. Likewise, Meyer-Mueller et al., (2020) found that VTEs are dependent on the dHPC, but not vHPC. Together, these findings suggest a strong link between VTEs and HPC-mPFC interactions. Our results are in agreement with this model, as we observed a significant reduction in VTE behavior following mPFC disruption. Furthermore, we found a significant correlation between the change in choice accuracy and the change in VTE occurrence due to mPFC disruption, suggesting that reduced deliberation may be one factor that influenced the observed impairments in SDA accuracy.

The results of our statistical analysis of changes in VTE occurrence between baseline and stimulation blocks lead us to conclude that mPFC disruption decreases the occurrence of VTEs, but that the effect is not epoch specific. This is in line with Papale et al.'s result (2016) that showed that sharp wave ripples at a reward location in previous trials were inversely related with the occurrence of VTEs on the current trial. In other words, they found that physiological events occurring outside of the choice epoch may still influence behaviors carried out as choices are made. While this interpretation fits our data well, it is worth noting that our results reveal a strong trend toward VTE occurrence decreasing more drastically when stimulation was applied during the choice epoch. This is consistent with our finding that choice accuracy was significantly lower during choice epoch stimulation as well as our finding that changes in VTE occurrence are correlated with changes in choice accuracy.

In addition to our main VTE findings, it should be noted that the SDA task used in this study may be uniquely suited, in contrast to other WM tasks, to study VTE behavior. This is due to our use of two opposed start arms that are pseudo-randomly selected at the end of each trial. Thus, in our SDA task there is an added element of uncertainty that requires animals to utilize an allocentric strategy that references landmarks in the testing room to determine which start arm the animal is currently located and which reward arm they traveled to previously. Use of an egocentric strategy would result in near chance performance in this task. In contrast, other versions of alternation testing have a single start arm. Thus use of either allocentric or egocentric strategy can be used to solve the task accurately. Reflective of this added uncertainty is our observation that during baseline trials, animals performed on average VTEs on 26.9% of trials, despite an overall high level of performance accuracy and being well trained on the SDA task. This stands in contrast to reports that well trained rats exhibit VTEs on a smaller number of trials (Schmidt et al., 2013). Future work should take the specific spatial strategy needed to solve different variations of SWM tasks into account, as tasks that can be solved with a pre-determined egocentric strategy may require less deliberation for high-levels of delayed alternation performance.

While VTEs have been found to occur in numerous WM tasks, our study confirms and adds to these findings by providing evidence that 1) VTE behavior is dependent on the mPFC in SWM tasks, 2) VTE behavior may be dependent on processes which occur throughout different epochs of SWM tasks (although to varying degrees), and 3) the reduction in VTE behavior may reflect a loss of cognitive deliberative processes which tends to impair SWM performance. These findings further implicate VTEs as the behavioral expression of deliberation which, in part, relies on intact mPFC function.

3.5 Conclusion

The findings of this study help to reconcile conflicting evidence regarding the mPFCs involvement in decision-making and WM processes by showing that the mPFC is not required to retain task-relevant information online over short delays during SWM tasks while additionally confirming the mPFCs role in processes occurring at decision points such as deliberation and choice. Lastly, our findings that a reduction in deliberative behavior correlated with impaired SWM performance further implicates VTEs as a reliable reflection of internal deliberation. Overall, our results fit with models that propose the mPFC is a crucial node in decision-making and WM networks during SWM tasks and that VTEs are a reflection of internal deliberative processes which also depend on the mPFC. Extending these models, we suggest that the mPFC dynamically interacts with the HPC memory system at specific task phases when rule information is required in order to deliberate and make a choice. Also, in tasks involving SWM, the mPFC is not required to retain information online during short delays.

References

Amemiya S, Redish AD. (2016). Manipulating decisiveness in decision making: effects of clonidine on hippocampal search strategies. J Neurosci 36: 814–827. doi:10.1523/JNEUROSCI.2595-15.2016.

Avigan PG, Cammack K, Shapiro ML (2020). Flexible spatial learning requires both the dorsal and ventral hippocampus and their functional interactions with the prefrontal cortex. Hippocampus. 1-12. DOI: 10.1002/hipo.23198

Baeg EH, Kim YB, Huh K, Mook-Jung I, Kim HT, Jung MW (2003). Dynamics of population code for working memory in the prefrontal cortex. Neuron 40(1):177-188.

Baker PM, Rao Y, Rivera ZMG, Garcia EM, Mizumori SJY (2019). Selective functional connectivity between the lateral habenula and hippocampus during different tests of response flexibility. Front Molec Neurosci 245:1-15. PMID:31680854

Benchenane K, Peyrache A, Khamassi M, Tierney PL, Gioanni Y, Battaglia FP, Wiener SI (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. Neuron 66:921-936.

Churchwell JC, Kesner RP (2011). Hippocampal–prefrontal dynamics in spatial working memory: interactions and independent parallel processing. Behav Brain Res 225:389–395.

Colgin LL (2013). Mechanisms and functions of theta rhythms. Ann Rev Neurosci 36: 295-312.

Dolleman-van der weel MJ, Griffin AL, Ito HT, Shapiro ML, Witter MP, Vertes RP, Allen TA (2019). The nucleus reunions of the thalamus sits at the nexus of a hippocampus and medial prefrontal cortex circuit enabling memory and behavior. Learn Mem 26: 191-205.

Fell J, Axmacher N (2011). The role of phase synchronization in memory processes. Nat Rev Neurosci 12:105–118.

Floresco SB, Seamans JK, Phillips AG (1997). Selective roles for hippocampal, prefrontal cortical, and ventral striatal circuits in radial-arm maze tasks with or without a delay. J Neurosci 17(5):1880-1890.

Goto Y, Grace AA (2007) Dopamine modulation of hippocampal-prefrontal cortical interaction drives memory-guided behavior. Cereb Cortex 18(6): 1407-14. doi:10.1093/cercor/bhm172

Guise KG, Shapiro ML (2017). Medial prefrontal cortex reduces memory interference by modifying hippocampal encoding. Neuron 94: 183-192.

Hallock HL, Wang A, Griffin AL (2016). Ventral midline thalamus is critical for hippocampal-prefrontal synchrony and spatial working memory. J Neurosci 36(32):8372-8389.

Horst NK, Laubach M (2012). Working with memory: evidence for a role for the medial prefrontal cortex in performance monitoring during spatial delayed alternation. J Neurophysiol 108(12): 3276-3288. doi: 10.1152/jn.01192.2011

Hyman JM, Zilli EA, Paley AM, Hasselmo ME (2010). Working memory performance correlates with prefrontal-hippocampal theta interactions but not with prefrontal neuron firing rates. Front Integr Neurosci 4:2.

Ito HT, Zhang SJ, Witter MP, Moser EI, Moser MB (2015). A prefrontal-thalamo-hippocampal circuit for goal directed spatial navigation. Nature 522: 50-55.

Jones MW, Wilson MA (2005). Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. PLOS Biol 3(12): 2187-2199.

Kesner RP, Churchwell JC (2011). An analysis of rat prefrontal cortex in mediating executive function. Neurobiol Learn Mem 96: 417–431. doi:10.1016/ j.nlm.2011.07.002.

Lee I, Kesner RP (2003). Time-dependent relationship between the dorsal hippocampus and the prefrontal cortex in spatial memory. J Neurosci 23:1517–1523

Maharjan DM, Dai YY, Glantz EH, Jadhav SP (2018). Disruption of dorsal hippocampal - prefrontal interactions using chemogenetic inactivation impairs spatial learning. Neurobiol Learn Mem 155: 351–360 doi:10. 1016/j.nlm.2018.08.023.

Meyer-Lindenberg AS, Olsen RK, Kohn PD, Brown T, Egan MF, Weinberger DR, Berman KF (2005). Regionally specific disturbance of dorsolateral prefrontal-hippocampal functional connectivity in schizophrenia. Arch Gen Psychiatry 62:379–386

Meyer-Mueller C, Jacob PYJ, Montenay JY, Poitreau J, Poucet B, Chaillan FA (2020). Dorsal, but not ventral, hippocampal inactivation alters deliberation in rats. Behavl Brain Res 390. DOI: https://doi.org/10.1016/j.bbr.2020.112622

Muenzinger KF (1938). Vicarious trial and error at a point of choice. I. A general survey of its relation to learning efficiency. J Genet Psychol 53: 75–86. doi:10.1080/08856559.1938.10533799

O'Neill PK, Gorgon JA, Sigurdsson T (2013). Theta oscillations in the medial prefrontal cortex are modulated by spatial working memory and synchronize the hippocampus through its ventral subregion. J Neurosci 33(35): 14211-14224.

Papale AE, Stott JJ, Powell NJ, Regier PS, Redish AD (2012). Interactions between deliberation and delay-discounting in rats. Cogn Affect Behav Neurosci 12: 513–526. doi:10.3758/s13415-012-0097-7.

Papale AE, Zielinski MC, Frank L, Jadhav SP, Redish AD (2016). Interplay between hippocampal sharp wave ripple events and vicarious trial and error behaviors in decision making. Neuron 92(5): 975-982.

Paz R, Bauer EP, Paré D. (2008). Theta synchronizes the activity of medial prefrontal neurons during learning. Learn Mem 15: 524–531. doi:10.1101/lm. 932408.

Pratt WE, Mizumori SJ (2001). Neurons in Rat Medial Prefrontal Cortex Show Anticipatory Rate Changes to Predictable Differential Rewards in a Spatial Memory Task. Behav Brain Res 123(2): 165-183.

Preston AR, Eichenbaum H. (2013). Interplay of hippocampus and prefrontal cortex in memory. Curr Biol 23: R764–R773. doi:10.1016/j.cub.2013.05. 041.

Redish AD (2016). Vicarious trial and error. Nat Rev 17: 147-159.

Santos-Pata D, Verschure PFMJ (2018). Human vicarious trial and error is predictive of spatial navigation performance. bioRxiv. doi: http://dx.doi.org/10.1101/332189.

Schmidt B, Duin AA, Redish AD (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. J Neurophysiol 121: 1981-2000.

Schmidt B, Papale A, Redish AD, Markus EJ. (2013). Conflict between place and response navigation strategies: effects on vicarious trial and error (VTE) behaviors. Learn Mem 20: 130–138. doi:10.1101/lm.028753.112.

Spellman T, Rigotti M, Ahmari SE, Fusi S, Gogos JA, Gordon JA (2015). Hippocampal–prefrontal input supports spatial encoding in working memory. Nature 522:309-314.

Tamura M, Spellman TJ, Rosen AM, Gogos JA, Gordon JA (2017). Hippocampal-prefrontal theta-gamma coupling during performance of a spatial working memory task. Nat Comm8: 2182.

Thierry AM, Gioanni Y, De´ge´ne´tais E, Glowinski J (2000). Hippocampo-prefrontal cortex pathway: Anatomical and electrophysiological characteristics. Hippocampus 10:411–419.

Warden MR, Miller EK (2010). Task-dependent changes in short-term memory in the prefrontal cortex. J Neurosci 30:15801–15810.

Xia M, Liu T, Bai W, Zheng X, Tian X (2019). Information transmission in HPC-PFC network for spatial working memory in rat. Behav Brain Res 356: 170-178.

Zielinski MC, Tang W, Jadhav SP (2017). The role of replay and theta sequences in mediating hippocampal prefrontal interactions for memory and cognition. Hippocampus 2020(30): 60-72.

Zylberberg J, Strowbridge BW (2017). Possible neural substrates for working memory. Ann Rev Neurosci 40:603-327.

Figure S-1. Rats' initial training performance on the spatial delayed alternation (SDA) task. Average choice accuracy (%, mean $\pm$ SD) is plotted on the y-axis, and the x-axis represents the training day. Rats reached criterion performance (dashed line) in about 7 days. They retained this high level of performance throughout training.

Figure S-2. In a separate group of animals we verified mPFC cellular responses to channelrhodopsin-2 light stimulation. Shown here are 4 different mPFC cells and their normalized firing rates in response to 50 and 100ms laser pulses which occurred 30 seconds apart. The vertical orange dotted lines represent the time laser onset and offset (Binwidths = 20ms). This diversity of artificially induced neural responses likely altered the spike timing relationships of neurons resulting in the functional disruption of the mPFC.

Figure S-3. Blue dots represent rat trajectories during the choice epoch of the plus maze (grey lines). On the left, 3 examples of rat trajectories characterized as vicarious trial and errors (VTE). On the right, 3 examples of rat trajectories characterized as non-VTEs. Note that all of the trajectories presented in this figure start at the south start arm, although the start arms on the Spatial Delayed Alternation (SDA) task pseudo-randomly alternated.

Chapter 4

The medial prefrontal cortex during flexible decisions: Evidence for its role in distinct working memory processes

A version of this chapter has been published as a pre-print on bioRxiv.

Citation:
Kidder, KS, Gillis, RM, Miles, JT, & Mizumori, SJY (2023). The medial prefrontal cortex during flexible decisions: Evidence for its role in distinct working memory processes. bioRxiv. https://doi.org/10.1101/2023.05.22.541807

Abstract:

During decisions that involve working memory, task-related information must be encoded, maintained across delays, and retrieved. Few studies have attempted to causally disambiguate how different brain structures contribute to each of these components of working memory. In the present study, we used transient optogenetic disruptions of rat medial prefrontal cortex (mPFC) during a serial spatial reversal learning (SSRL) task to test its role in these specific working memory processes. By analyzing numerous performance metrics, we found: 1) mPFC disruption impaired performance during only the choice epoch of initial discrimination learning of the SSRL task, 2) mPFC disruption impaired performance in dissociable ways across all task epochs (delay, choice, return) during flexible decision-making, 3) mPFC disruption resulted in a reduction of the typical vicarious-trial-and-error (VTE) rate modulation that was related to changes in task demands. Taken together, these findings suggest that the mPFC plays an outsized role in working memory retrieval, becomes involved in encoding and maintenance when recent memories conflict with task demands, and enables animals to flexibly utilize working memory to update behavior as environments change.

4.1 Background and Introduction

Memory-guided decisions are successful when prior learning matches our current expectations. When goals change, however, we must flexibly update both our behavior and the memories we use to guide our decisions. This process of updating is said to rely on working memory which involves encoding of new information, as well as maintenance and retrieval of that information to guide upcoming behavior (Baddeley, 2011; Becker et al., 1981; Buzsáki et al., 2021; Cohen et al., 1997).

Neurophysiology studies have found correlates for each of these three components of working memory (encoding, maintenance, retrieval) in prefrontal cortices. Indicative of encoding, medial prefrontal cortex (mPFC) cells respond to choice outcomes (Horst & Laubach, 2012; Luk & Wallis, 2009; Pratt & Mizumori, 2001; Y. Yang & Mailman, 2018) in tasks that require use of different strategies to earn reward. mPFC cells can also encode switches in strategy use (Hasz & Redish, 2020) and generalized task variables (Samborska et al., 2022). Together, these studies suggest that the mPFC can represent information about rewards, task structures, and rules that can be encoded for later retrieval to guide behavior.

Cells in the primate prefrontal cortex (PFC) can also exhibit elevated activity throughout the duration of a delay period, which is when information needs to be maintained in working memory to guide an upcoming decision (Batuev et al., 1979; Funahashi et al., 1989; Fuster & Alexander, 1971; Kubota & Niki, 1971). Similar delay-firing has been found in the rodent mPFC along with populations of cells whose collective activity tiles delay periods (Baeg et al., 2003; Bolkan et al., 2017; Jung et al., 1998). These results have often been interpreted as a physiological basis for active working memory maintenance (Goldman-Rakic, 1995; Zylberberg & Strowbridge, 2017).

In addition to working memory encoding and maintenance, mPFC cells have been shown to retrieve task-relevant information that guides current and upcoming choices. These cells can fire specifically at or before choice points where rats turn one direction or another in spatial delayed alternation tasks, and their firing can predict success on these tasks (Guise & Shapiro, 2017; Ito et al., 2015; Luk & Wallis, 2009; Stout & Griffin, 2020; S. T. Yang et al., 2014; Y. Yang & Mailman, 2018). Many studies also show increased hippocampal (HPC)-mPFC oscillatory coherence in the 4-12 Hz theta band surrounding choice points (Benchenane et al., 2010; Jones & Wilson, 2005; Tamura et al., 2017), suggesting choice points of tasks are times when information is retrieved and potentially shared across structures to guide decision-making.

Lending support to physiological studies, behavioral studies involving manipulations of the mPFC also suggest that the mPFC plays a role in working memory processing. Lesions and pharmacological inactivations decrease choice accuracy in spatial delayed alternation and non-match to place tasks, and impair switches between reward contingencies in reversal learning and set shifting tasks demonstrating impairments in working memory (Avigan et al., 2020; Birrell & Brown, 2000; De Bruin et al., 2000; Kinoshita et al., 2008; Ragozzino et al., 1999). These studies, however, were unable to address whether distinct components of working memory (i.e. encoding, retrieval, maintenance) were affected by the mPFC manipulations because of their permanent or long-lasting effects.

Recent work using temporally precise mPFC manipulations has begun to assess causal links between disrupted mPFC function and working memory processes. Specifically, inhibiting somatostatin (SST) or parvalbumin (PV) expressing cells in the mPFC during early delay periods impaired performance on an auditory go/no-go task (Kamigaki & Dan, 2017). Another study found that either activating or suppressing mPFC activity during delays in an odor-based non-match to sample task decreased task performance when mice were learning a discrimination rule (Liu et al., 2014). Interestingly, well trained mice performed worse when mPFC activity was suppressed during the decision-making/sample matching portion of the task, but not when the mPFC was manipulated during the delay (Liu et al., 2014), suggesting the mPFC is involved in distinct working memory processes as behavior is shaped by learning.

We followed up on these reports by disrupting the mPFC at different epochs of a decision in a spatial delayed alternation (SDA) task (Kidder et al., 2021). Starting with the premise that different working memory processes were associated with different stages of the task, we reasoned that encoding began during reward delivery as rats returned to a waiting location (the return epoch), working memory maintenance occurred as rats held information about their prior choice during a delay period (delay epoch), and retrieval into working memory occurred as rats made choices about where to go for reward (choice epoch). Thus, decision-making impairments could be associated with different working memory processes based on the epoch when mPFC disruption occurred. Interestingly, we found deficits in choice accuracy when the mPFC was disrupted only during the choice epoch, but not delay or return epochs, suggesting a major role for the mPFC in working memory retrieval during spatial delayed alternation. Additionally, disrupting the mPFC during any epoch caused a decrease in a decision-making behavior known as vicarious trial and error (VTE), where decision-makers appear to vacillate between options before settling on a final decision. Furthermore, the magnitude of the decreased VTE occurrence was correlated with the magnitude of the choice accuracy deficit.

However, because the SDA task only required encoding which location was visited every trial, accurate decision-making did not depend on encoding or maintaining reward information about recent decisions. To ensure that successful performance relied on encoding and maintaining information about task-relevant variables, in this study we employed the same epoch-based disruption procedure on a serial, spatial reversal learning (SSRL) task, where reward locations switched (reversed) dynamically based on recent performance. If, as in Kidder *et al.* (2021), we assume the working memory processes are reliably separated throughout trial epochs, this design allows us to make predictions about what behavior should look like when different working memory processes are disrupted. Much like the SDA task, disrupting the mPFC during retrieval should always cause choice accuracy impairments, leading to fewer reversals and increases in the average number of trials to criterion. Because new reward contingencies need to be encoded as reversals occur, disrupting the mPFC during encoding should lead to errors caused by regression to behavior consistent with the prior contingency. Similarly, accurate decision-making should rely on maintaining information about prior choices until new contingencies are re-encoded. This implies that disrupting the mPFC during working memory maintenance should also cause performance deficits on trials close to reversals, leading to perseveration of behavior consistent with the prior contingency.

Our results indeed show that disrupting the mPFC during each epoch can cause performance deficits once rats are required to update their behavior. The most severe deficits were caused by disruption during the choice epoch, suggesting that the mPFC plays an outsized role in retrieving information into working memory. Furthermore, disruption during the delay epoch caused increases in the number of errors immediately following reversals without altering other performance metrics, and disruption during the return epoch caused fewer total reversals per session. What's more, mPFC disruption, regardless of epoch, prevented the typical aggregation of vicarious trial and error behavior around reversals without decreasing their overall prevalence, suggesting a general failure to correctly time flexible behavior with respect to changes in task demands.

4.2 Methods

4.2.1 Animals

Five Long-Evans rats (Charles River) were used in this study. The cohort consisted of 3 males (320-400 grams) and 2 females (180-220 grams). Animals were housed on a 12 hour light/dark cycle (lights on at 7:00 am) with *ad libitum* access to water. Rats were free fed upon arrival for one week, after which they were food restricted to 80-85% of their original free fed weight. Animals were only trained and tested during the light portion of their light/dark cycle. All procedures were in accordance with the University of Washington's Institutional Animal Care and Use Committee guidelines (Protocol 3279-01).

4.2.2 Apparatus

The spatial serial reversal learning (SSRL) task took place on a fully automated plus-maze elevated 79 cm from the floor. Arms of the maze measured 58 x 5.5cm. The north and south arms were designated as start-arms, and the east and west arms were designated as goal-arms. Attached to the end of the goal-arms were 3D printed food wells connected to computer-controlled pellet delivery hardware (Med-Associates Inc.) which delivered sucrose pellets (45mg; TestDiet). The maze was remotely controlled by LabVIEW 2016 software (National Instruments) with custom built task programs. Each maze arm was hinged midway so that the proximal end could be raised and lowered via servos connected to Arduino boards. The maze was surrounded by black curtains with several visual cues attached to them so that animals could use these cues to engage spatial navigation strategies. Positioned directly above the maze was a camera (SONY) recording at ~30Hz which integrated with LabVIEW software to identify animal location and trigger task events based on the coordinates of predetermined trigger locations.

4.2.3 Surgical Procedures and Optic Implants

Shortly after arriving at our facility, rats were anesthetized using 5% isoflurane in oxygen (flow rate 1.0 L/min) and placed into a stereotaxic apparatus (KOPF). Isoflurane concentration was then lowered to 1.0%-3.5% as rats underwent surgery involving bilateral mPFC (AP: 3.0mm, ML: ± 0.8, DV: -3.8) intracranial injections of the excitatory optogenetic viral construct AAV5-CaMKlla-hChR2-mCherry (Addgene: CS1096). 500nL of virus was injected into each hemisphere at a flow rate of 100nL/min. Following surgery, rats were allowed approximately seven days of recovery before beginning handling and maze training procedures.

Once animals reached performance criterion on the SSRL task, they underwent optic fiber implant surgery. Prior to implant surgery, optic implants were constructed using optic fiber (200μm in diameter) and ceramic ferrules held together with a quick-cure epoxy (ThorLabs). Custom implant devices which housed the two optic fibers were designed (Autodesk Inventor) and then 3D printed (Formlabs). For implant surgery, isoflurane conditions were the same as previously described, and holes were drilled into the skull at the previously used bilateral mPFC coordinates. Tips of the optic fibers were positioned just above the location of the previous viral injections (~D/V:-3.5). Approximately 6-8 stainless steel screws were anchored into the skull, and dental repair resin (Coltene) was applied to cover the screws and hold the entire optic implant device in place. Animals were once again given approximately seven days to recover from surgery. After recovery, animals were required to meet performance criterion three consecutive days before experimental conditions began.

4.2.4 Behavioral Training and Experimental Design

Habituation and Training

Rats were handled for 10-15 minutes on at least three occasions before they were exposed to the maze. Rats were habituated to the maze prior to behavioral training by allowing them to freely forage for sucrose pellets scattered on the maze for one session of 20 minutes. Next, animals performed a training program which consisted of 45 trials and had alternating blocks of forced choice and free choice trials in which every response was rewarded. Animals were required to complete the training program within 45 minutes for three days in a row before moving on to training on the SSRL task.

Spatial Serial Reversal Learning (SSRL) Task

The goal of the SSRL task was for animals to disambiguate which of the two goal locations (east or west) was the current block's correct (reward) location. At the start of each session the initial correct arm was randomly set by the experimenter and each block ended when the animal chose the correct arm 9/10 times. At the beginning of each new block the reward location was switched to the opposite reward arm (**Figure 4.1a**). Animals were then tested to see how many reward arm reversals they could complete within a session. The first block of a session was considered the initial discrimination (ID) block, and all subsequent blocks were considered reversal blocks.

A single session of the SSRL task consisted of 200 trials and each trial consisted of three epochs: delay, choice, and return (**Figure 4.1A,B**). At the start of each trial, all maze arms were lowered so that the animal was restricted to their current start-arm where they waited through the five second delay epoch. After the delay, the current trial's start-arm and both goal-arms were raised so that the animal could navigate to the goal location of their choice. A typical choice epoch, which started by raising the start arm after the delay and ended when an animal reached a goal location, lasted 1.5-3 seconds. After reaching a goal location, animals received one sucrose pellet for correct responses and no sucrose pellets for incorrect responses. During this time, the next trial's start-arm was randomly chosen and subsequently raised so that the animal could travel there and trigger the start of the next trial's delay. Therefore, the return epoch was defined from the point of the animal reaching a goal location, included consuming (or not consuming reward), and ended when the animal reached the next trial's randomly chosen start-arm. Return epochs generally lasted five seconds. Animals reached performance criterion on the SSRL task once they reached asymptotic performance which was measured as the total number of reversals completed in a session. After reaching the initial SSRL performance criterion animals underwent optic fiber implantation surgery.

Experimental Design

Each animal underwent a total of five experimental conditions which occurred in separate sessions, each on a separate day (**Figure 4.1c**). These conditions included: (1) a pre-stimulation assessment, (2-4) mPFC optogenetic disruption during each of the three task epochs (delay, choice, return; order of epoch-specific disruption sessions was counterbalanced for each animal), and (5) a post-stimulation assessment. There were no significant differences between pre- and post-stimulation assessments, indicating no lasting effects of optogenetic stimulation (Figure S1). Therefore, the average of pre- and post-stimulation assessment sessions were combined to form the baseline for comparison in subsequent data analyses.

Optogenetic Stimulation

To activate opsins and disrupt mPFC neural activity, a 473nm laser (Laserglow Technologies) was held to a power of ~6-7mw emanating from the tip of the implanted optic fibers (measured prior to optic implant surgery). Stimulation was delivered at a rate of 20 Hz with a 50% duty cycle for the duration of the selected epoch (delay, choice, or return). Stimulation rate was controlled by an Arduino which also received signals from the LabVIEW program to gate laser activation according to the specific epoch being tested. A single fiber optic cable was connected from the power source to an optic commutator (Doric) mounted above the maze, from which two fiber optic cables were affixed. These fiber optic cables were then connected to the animal's optic ferrules with copper sleeve connectors (ThorLabs).

4.2.5 Histology

After all conditions were completed, rats were given an overdose of sodium pentobarbital. Once rats were deeply anesthetized, they were transcardially perfused with 0.9% saline and 10% formaldehyde solution. Brains were extracted and stored at 4°C in formalin for a day and then submerged in a 30% sucrose solution for four days. Brains were then frozen and cut into coronal sections (45 µm) on a freezing microtome. Brain slices were then mounted onto slides and fluorescence was preserved with the mounting medium Vectashield (Vector Laboratories). Slices were examined with a fluorescent microscope to verify viral expression and optic fiber placement in the mPFC. Only animals with proper bilateral viral expression and optic fiber placement contributed to subsequent data analysis.

4.2.6 VTE Identification and Analysis

Using the same videos recorded during in-task position tracking (see **Apparatus** above), we used DeepLabCut (DLC) version 2.2 to identify the heads of implanted rats running the SSRL task. We initially labeled 20 or 30 frames for each session using the built-in labeling GUI for a total of approximately 400 labeled frames for the first model training attempt. Because many of our images were low contrast and mislabeled, we relabeled 20-50 outliers in a subset of videos and retrained a new iteration of the network on the now expanded dataset to achieve satisfactory performance, confirmed by manual inspection of trajectories. Each training attempt used NVIDIA GEFORCE GTX 1080 GPU with 500,000 iterations.

Choice epoch segmentation for trajectory detection was accomplished by detecting when rats crossed the maze central platform, then moving earlier in the trajectory to a user-defined starting point and later in the trajectory to a user-defined ending point. Before epoch segmentation, all position data from DLC were rotated and scaled into maze coordinates with the center of the maze at approximately the origin of an (x, y) grid. Position data were median filtered with a 7-point window to mitigate any jumps from DLC tracking and all choice epoch trajectories were quality checked by experimenters.

Trajectories with vicarious trial and error (VTE) were detected by projecting the position data into principal component (PC) space and clustering the PC-representations of the trajectories with hierarchical agglomerative clustering. Before projection, all trajectories were aligned and standardized to the same starting and ending positions. Visual inspection of clustering in PC space naturally formed what looked like two clouds in low dimensional plots, and distance-based dendrograms cut to give two clusters separated trajectories with VTE from non-VTE trajectories.

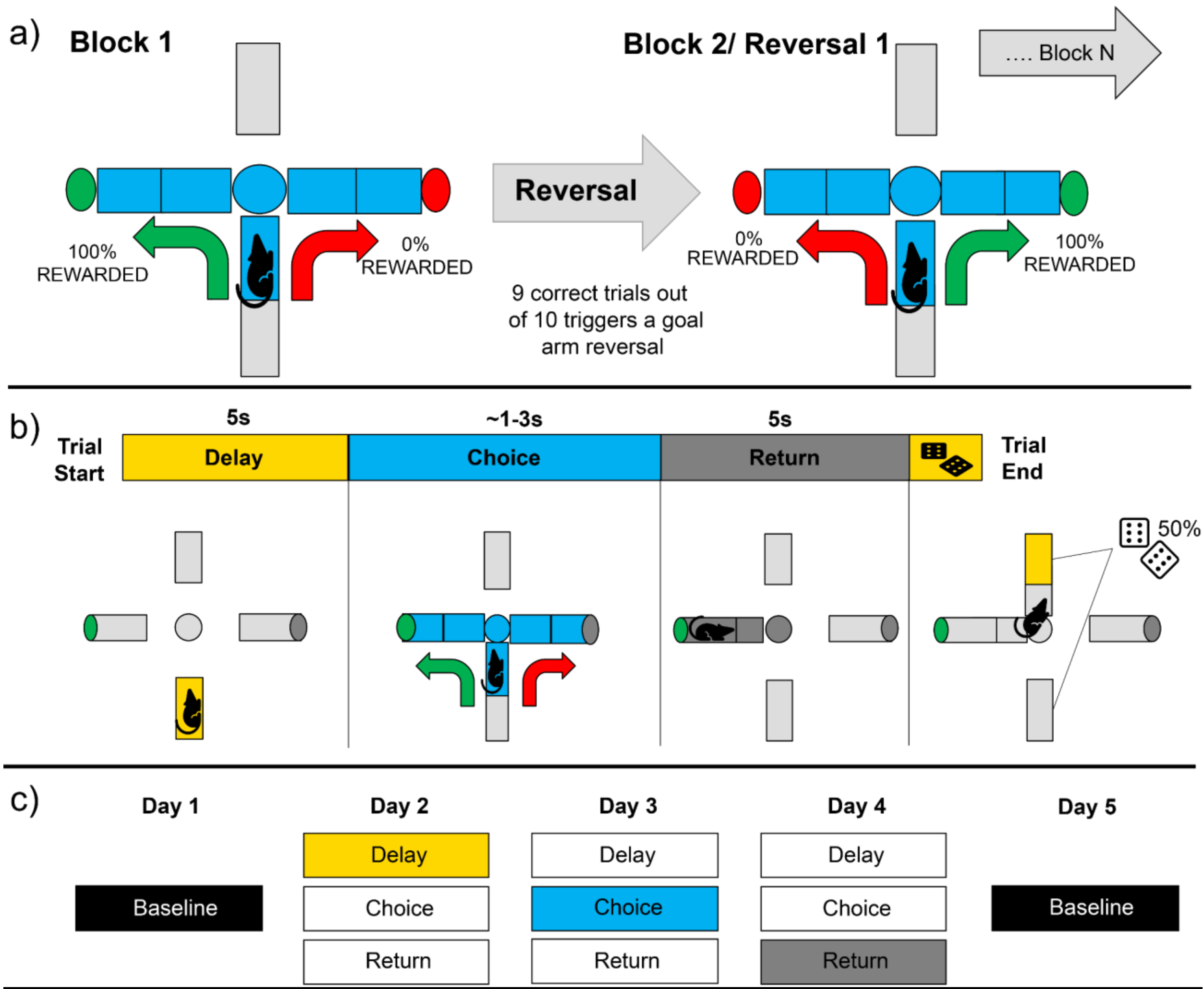

Figure 4.1: (a) Animals first showed discrimination of the initially rewarded arm by selecting the correct arm 9 out of 10 times. After this, the reward location was reversed to the opposite reward arm. **(b)** Depiction of SSRL task epochs (delay, choice, return). **(c)** Animals

underwent five days of experimental sessions. Animals started with a baseline session when no optogenetic disruption occurred, then in the subsequent three days the animal experienced optogenetic mPFC disruption in a selected epoch for every trial in the session (stimulated epoch order was randomized across animals). Experimental sessions concluded with a final baseline day where no mPFC disruption occurred.

4.3 Results

4.3.1 Histology

Tips of optic fibers were located bilaterally in the mPFC (**Figure 4.2a**). Optic fiber tip placements were evenly distributed along the D/V axis of the prelimbic cortex, ranging from just below the ACC-PL border to the PL-IL border. Viral expression in the mPFC surrounded all optic fiber tips, and expression was relatively equal between mPFC hemispheres. Animals without proper optic fiber placement or viral expression were not included in the data analysis.

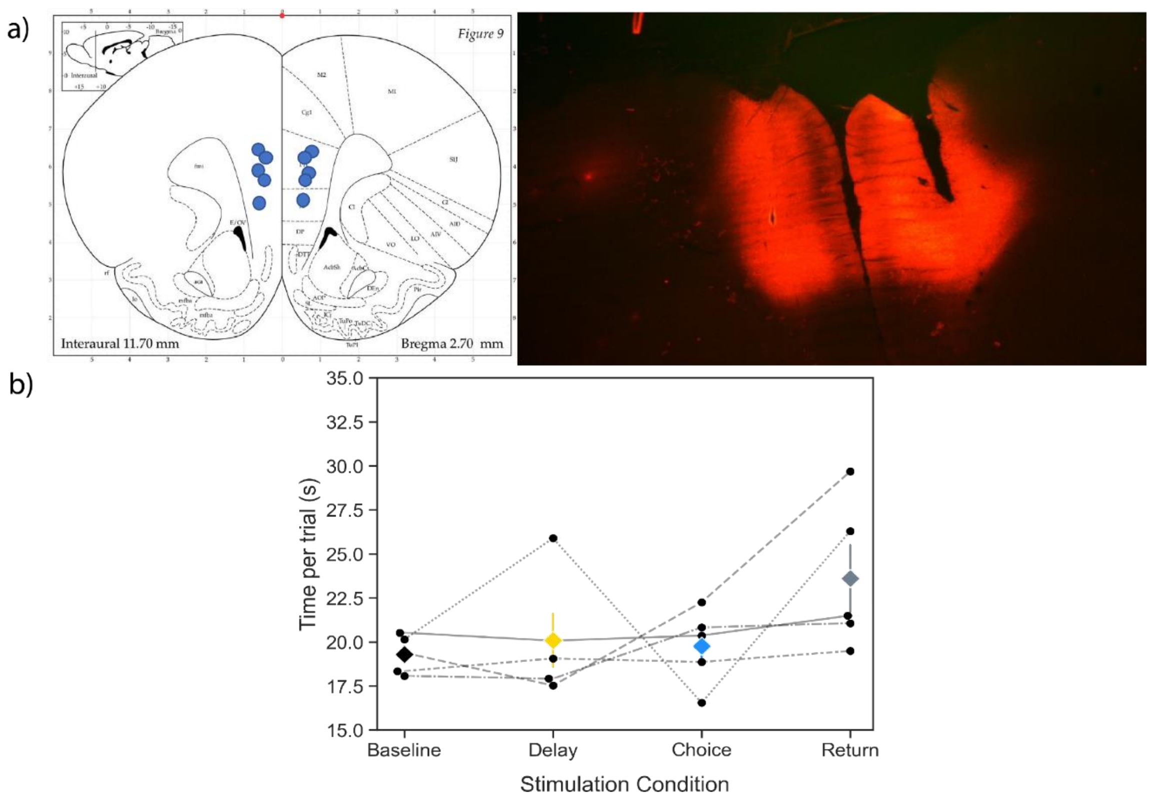

Figure 4.2: (a) mPFC optic fiber tips terminated within the prelimbic/infralimbic division of the mPFC. Only animals with clear bilateral mPFC expression were used for analysis. **(b)** Time per trial was not significantly different ($p > .05$) between conditions, suggesting stimulation did not impair animals' motivation to complete the task.

4.3.2 mPFC disruption impairs SSRL performance

To validate that the mPFC is involved in this task, we combined all stimulation conditions together and compared the distributions of average trials per block for stimulation ($\bar{x} = 18.8$, SE = 1.05) and baseline conditions ($\bar{x}$ = 13.9, SE = 0.47) (**Figure 4.3a**). A t-test revealed the distributions are significantly different from each other (t = -6.89, $p < 0.05$). These data reveal that regardless of which epoch disruption was applied to, mPFC disruption impaired performance on the SSRL task. Additionally, time per trial was not significantly different between baseline or stimulation conditions (t = -1.12 , $p = 0.277$) suggesting mPFC disruption did not alter animals' motivation to engage in the task (**Figure 4.2b**).

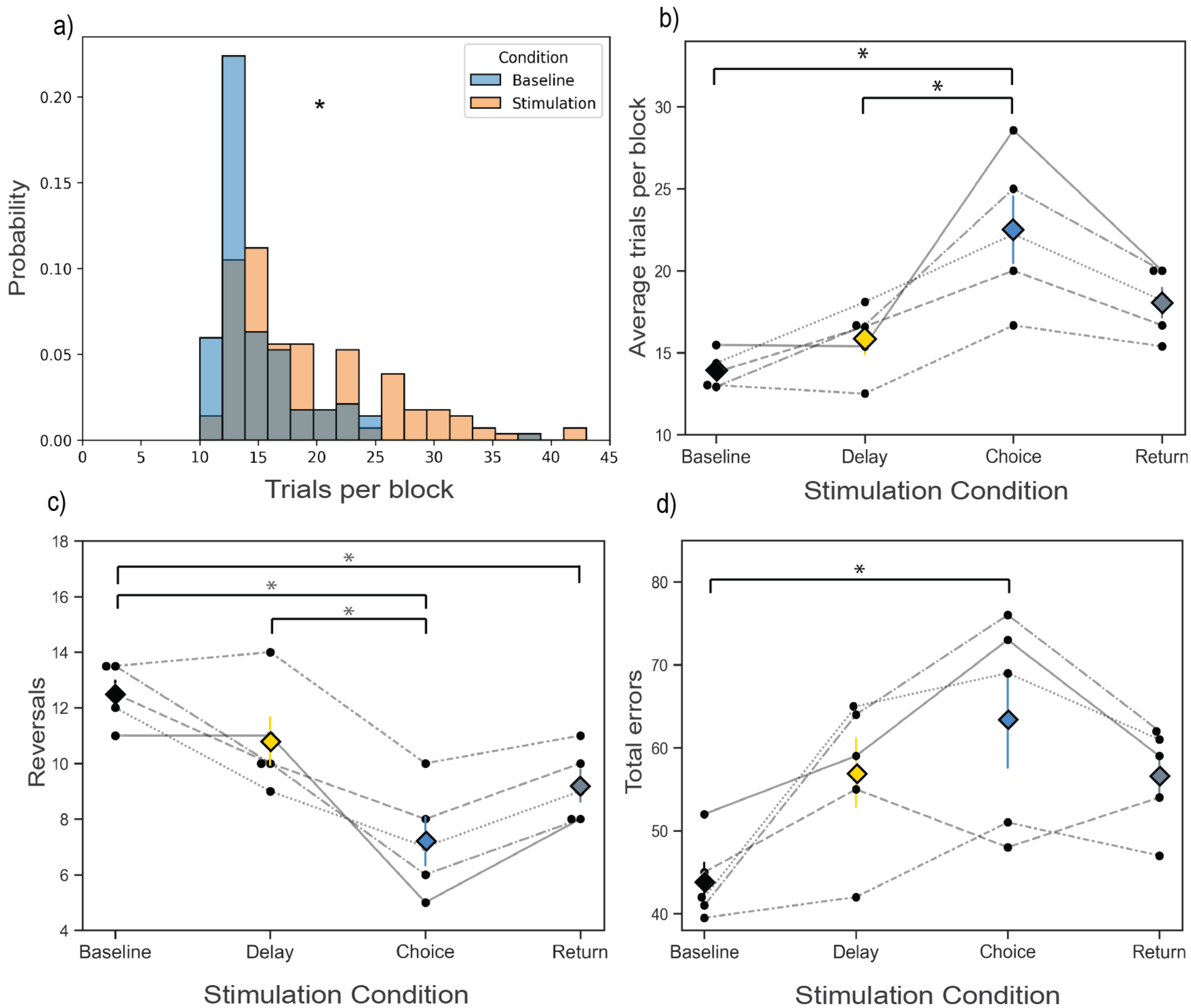

Figure 4.3 (a) There was a significant difference (*$p < 0.05$) in the probability of trials per block between baseline and stimulated sessions. **(b)** Choice epoch disruption significantly increased the trials in a block compared to delay epoch disruption and baseline. **(c)** Choice epoch disruption significantly reduced the number of reversals completed in a session compared to delay epoch disruption, and choice and return epoch disruption significantly reduced the number of reversals completed in a session compared to baseline. **(c)** Choice epoch disruption significantly increased the total number of errors compared to baseline.

4.3.3 Differential epoch-specific effects on SSRL performance

To determine if mPFC disruption during a particular epoch (delay, choice, return) selectively impaired performance, we compared the average trials per block, average number of reversals, and total errors across baseline and the three epoch stimulation conditions (**Figure 4.3b-d**). There was an effect of stimulation epoch on the average trials per block ($F(3) = 15.67$, $p = 0.0002$) in which choice epoch stimulation caused significantly more trials per block ($\bar{x} = 22.49$, SE = 2.04) than baseline ($\bar{x} = 13.92$, SE = 0.47) and delay stimulation ($\bar{x} = 15.85$, SE = 0.94), but not return stimulation ($\bar{x} = 18.04$, SE = 0.91) (**Figure 4.3b**). A repeated measures ANOVA revealed an effect of stimulation epoch on the average number of reversals ($F(3) = 23.93$, $p < 0.0001$). Post-hoc tests for multiple comparisons revealed the average number of reversals in the choice epoch stimulation condition ($\bar{x} = 7.20$, SE = 0.86) was significantly less than baseline ($\bar{x} = 12.5$, SE = 0.47) and delay epoch stimulation ($\bar{x} = 10.80$, SE = 0.86), but was not different from return stimulation ($\bar{x} = 9.20$, SE = 0.58), which also showed significantly less reversals than baseline (**Figure 4.3c**). There was a significant stimulation epoch effect on the total number of errors ($F(3) = 9.02$, $p = 0.0021$): stimulating during the choice epoch caused significantly more errors ($\bar{x} = 63.4$, SE = 5.8) than baseline ($\bar{x} = 43.9$, SE = 2.21), but not delay ($\bar{x} = 57.0$, SE = 4.16) or return stimulation ($\bar{x} = 56.6$, SE = 2.77) (**Figure 4.3d**). To summarize, stimulation across epochs tended to impair performance. However, choice epoch stimulation was the only epoch in which performance was consistently and significantly impaired. Return epoch stimulation significantly impaired performance only by reducing the total numbers of reversals in a session, while delay epoch stimulation did not significantly impair performance on any of these metrics.

4.4 mPFC choice epoch disruption impairs both initial discrimination and reversal performance

While there is consensus regarding the mPFC's involvement in discriminating between similar recent memory representations (goal locations) during reversal blocks of similar tasks, prior work (Avigan et al., 2020; Guise & Shapiro, 2017) has shown conflicting results regarding the mPFCs role in initial discrimination learning. We therefore analyzed the effects of mPFC disruption between ID and reversal blocks of this task (**Figure 4.4**). There was an effect of stimulation epoch on both the number of trials to ID ($F(3) = 3.97$, $p = 0.035$) and trials per reversal ($F(3)= 17.42$, $p = 0.0001$). In both cases, choice epoch stimulation increased the average number of trials (ID: $\bar{x} = 25.0$, SE = 2.98; reversal: $\bar{x} = 27.41$, SE = 1.23) compared to baseline (ID: $\bar{x} = 15.80$, SE = 0.93; reversal: $\bar{x} = 15.50$, SE = 0.26) (**Figure 4.4a-b**). We next analyzed the occurrence of errors between ID and reversal blocks. There was an effect of stimulation epoch on both the number of ID errors ($F(3) = 3.35$, $p = 0.056$) and average reversal errors ($F(3) = 11.39$, $p = 0.0008$) (**Figure 4.4c-d**). In both cases, choice epoch stimulation increased the number of errors committed (ID: $\bar{x} = 6.6$, SE = 1.21, reversal: $\bar{x} = 7.62$, SE = 1.23) compared to baseline (ID: $\bar{x} = 2.8$, SE = 0.34, reversal: $\bar{x} = 3.12$, SE = 0.26). Lastly, repeated measures ANOVAs revealed no effect of stimulation condition on ID accuracy ($F(3) = 1.52$, $p = 0.26$) but did reveal an effect on reversal accuracy ($F(3) = 9.40$, $p = 0.0018$) (**Figure 4.4e-f**). Choice epoch stimulation decreased reversal accuracy ($\bar{x} = 0.59$, SE = 0.033) compared to baseline reversal accuracy ($\bar{x} = 0.72$, SE = 0.014) but was not different from return ($\bar{x} = 0.64$, SE = 0.016) or delay ($\bar{x} = 0.65$, SE = 0.021) which were both not significantly different from baseline.

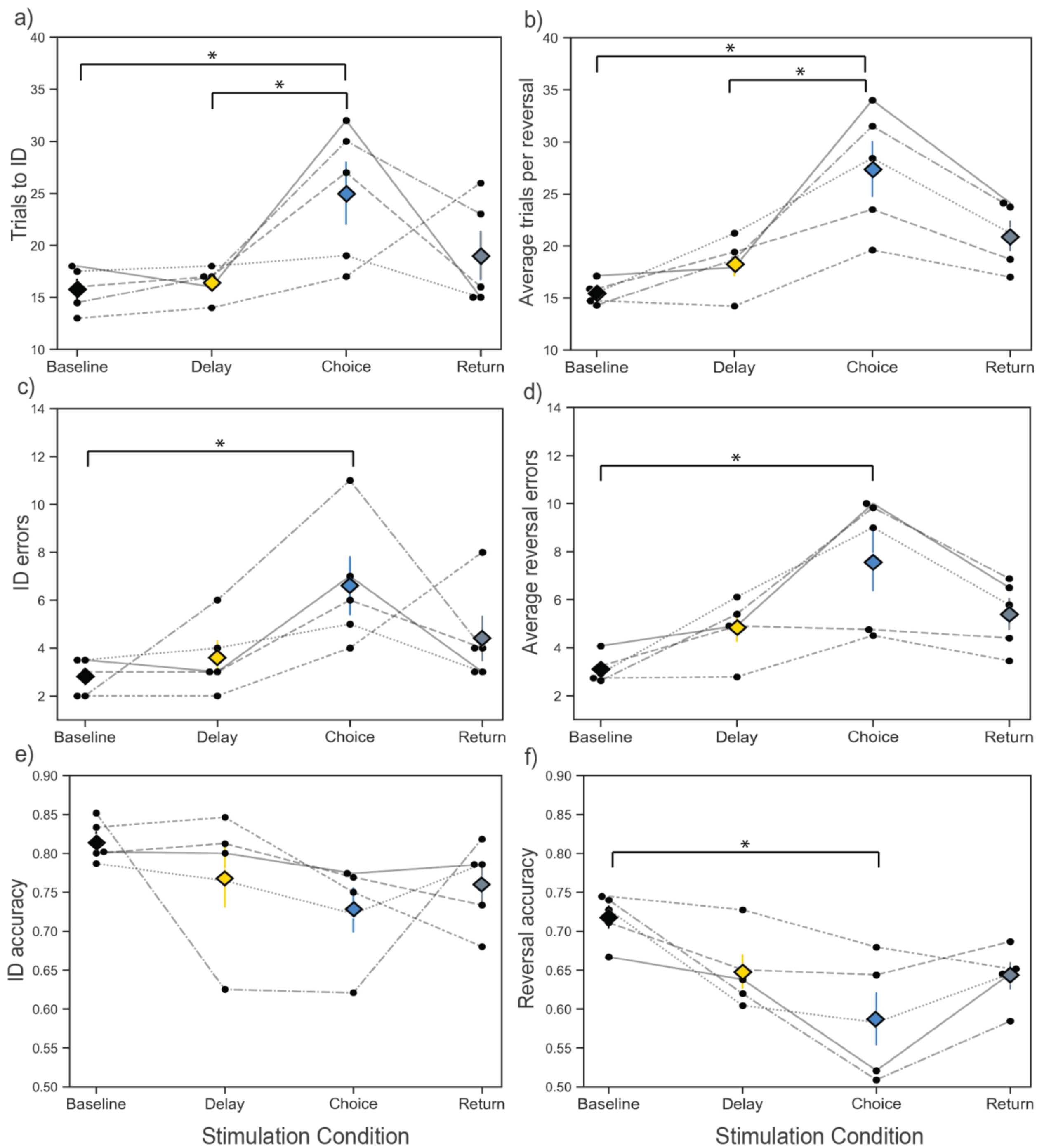

Figure 4.4 Choice epoch disruption significantly increased (*$p < 0.05$) both the number of trials to initial discrimination (ID) **(a)** and the average trials per reversal **(b)**. Choice epoch disruption significantly impaired the number of errors in ID **(c)** and reversal **(d)** blocks. ID choice accuracy **(e)** was not impaired by mPFC disruption in any condition, however choice disruption did significantly impair reversal block choice accuracy **(f)**.

These results reveal mPFC stimulation impaired both ID and reversal performance to some extent. However, while impacts to ID and reversal performance were similar across most metrics, reversal performance was impacted to a greater extent as revealed by a selective decrease in choice accuracy in reversal blocks and not ID blocks following choice epoch disruption.

4.3.5 Perseverative and regressive errors reveal dissociable working memory roles for the mPFC

We identified perseverative errors as all errors that occurred before the animal made two correct choices in a new block, while regressive errors were those that occurred after the animal made two correct choices in the new block. There was an effect of stimulation epoch on both the number of perseverative ($F(3) = 9.48$, $p = 0.0017$) and regressive ($F(3) = 8.59$, $p = 0.0026$) errors (**Figure 4.5a-b**). Animals committed significantly more perseverative errors during the delay epoch stimulation condition ($\bar{x} = 1.97$, SE = 0.26), compared to baseline ($\bar{x} = 1.09$, SE = 0.14), while errors following choice ($\bar{x} = 1.89$, SE = 0.22) and return ($\bar{x} = 1.54$, SE = 0.21) stimulation were not significantly different from baseline. Conversely, there were significantly more regressive errors in the choice epoch condition ($\bar{x} = 3.97$, SE = 0.93) than baseline ($\bar{x} = 1.02$, SE = 0.18) and delay ($\bar{x} = 1.50$, SE = 0.31) conditions, but not return condition ($\bar{x} = 2.40$, SE = 0.30).

To see how regressive and perseverative errors were distributed in individual blocks, we created an error index metric which measures ratio of perseverative to regressive errors, which allowed us to ask if this ratio differs by stimulation condition. If more regressive errors than perseverative errors occurred in a block, the error index value is positive. The error index value for a block is negative if more perseverative errors than regressive errors occurred in that block. During the ID there were no significant differences between the error index values of stimulation epoch conditions ($F(3) = 0.11$, $p = 0.96$) (**Figure 4.5c**). Furthermore, all stimulation conditions contained slightly positive error index values, indicating there were always more regressive errors during ID blocks. During reversal blocks ($F(3) = 4.71$, $p = 0.02$) however, delay epoch stimulation resulted in a switch from a positive to a negative error index value, and this was significantly different from the choice epoch stimulation condition ($p = 0.007$) (**Figure 4.5d**). These data reveal that mPFC disruption during reversal blocks, not ID blocks, biased the types of errors that occur in a block depending on which epoch the mPFC disruption occurred.

4.3.6 VTE rate fluctuates according to task demands and is impaired by mPFC disruption

Out of the 3561 trials analyzed, 686 (19.3%) were identified as VTE trials. Unlike Kidder et al. (2021), we do not see differences in VTE rates per block when we disrupt the mPFC ($t = 0.621$, $p = 0.535$, **Figure 4.6b**). Instead, mPFC stimulation altered the dynamics of VTE occurrence surrounding reversals (**Figure 4.6a,b** center and right columns). During baseline sessions, the average rate of VTE occurrence is higher than average early in the block, while near the end of a block the rate of VTE occurrence is lower than average (**Figure 4.6b**, top-center). Furthermore, the rates appear periodic, showing regularly spaced peaks in the average reversal-aligned VTE-rate autocorrelation functions (**Figure 4.6b**, top-right). In contrast, sessions with stimulation do not tend to show the same regular increases and decreases in VTE rate aligned to reversals. Rather, these sessions tend to show the same average VTE rate (approximately 20% of trials) regardless of proximity to a reversal. This is further indicated by the almost entirely flat autocorrelation functions for reversal-aligned VTE rates from stimulation sessions.

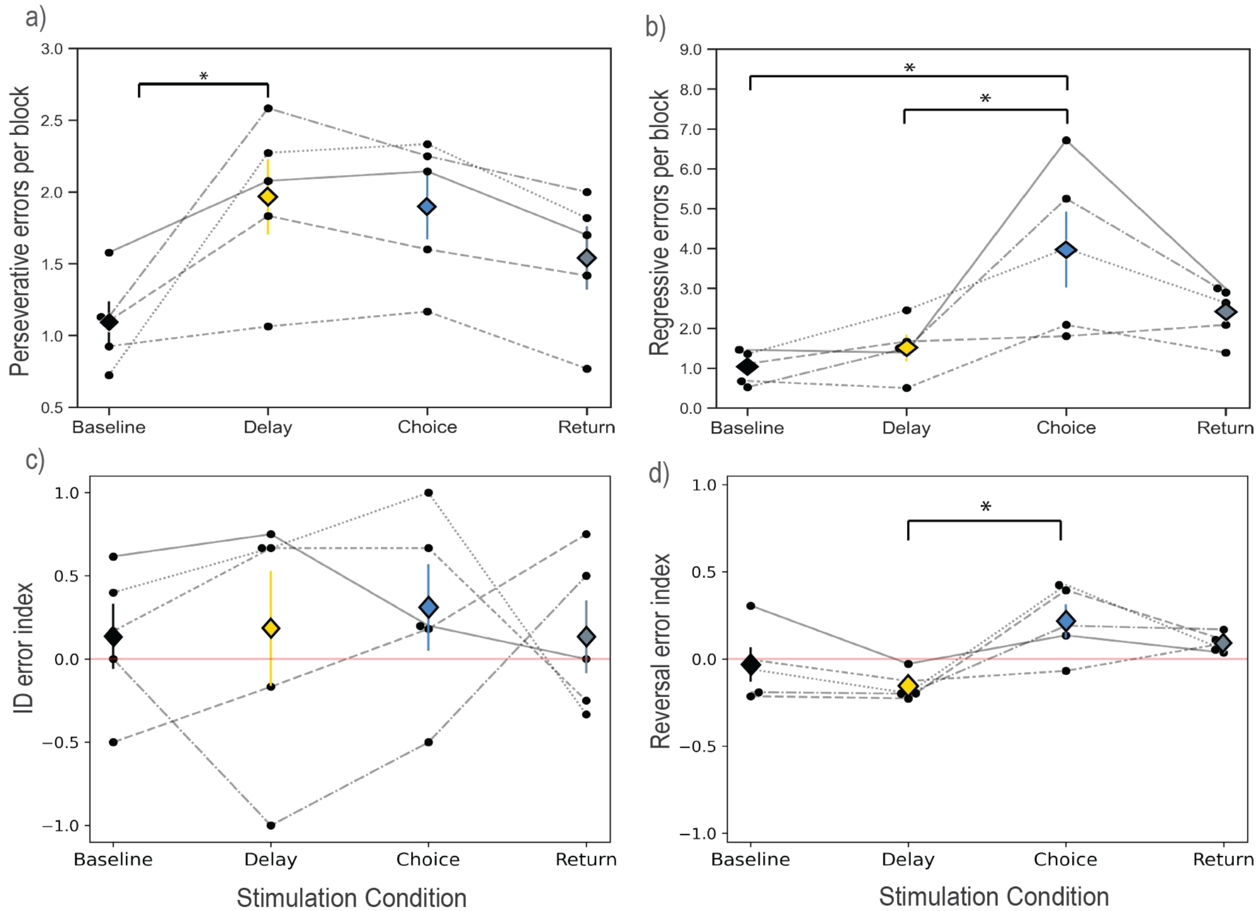

Figure 4.5 (a) Delay epoch disruption significantly increased (*p < 0.05) perseverative errors compared to baseline. After a reversal, perseverative errors were defined as all errors that occur before the animal makes 2 correct choices in a row. **(b)** Choice epoch disruption significantly increased regressive errors above baseline and delay epoch disruption. After the animal makes two correct choices in a row all subsequent errors in the block are defined as regressive errors. **(c)** There were no significant differences between average error index values by stimulation condition in the ID block. A positive error index means that more regressive errors than perseverative errors occurred in a block, while a negative error index means that more perseverative errors occurred than regressive errors in a block. **(d)** Average error index values in reversal blocks for the delay and choice disruption groups were significantly different.

4.4 Discussion

Findings from this study provide direct causal evidence of mPFC involvement, albeit in different ways, in distinct working memory processes during flexible decision-making in the SSRL task. Consistent with past studies, we found that bilateral mPFC optogenetic disruption impaired overall performance on the SSRL task. Furthermore, selectively, and separately disrupting the mPFC in each task epoch (delay, choice, return) impaired performance during reversal blocks, while mPFC disruption during initial discrimination (ID) blocks impaired performance only when the disruption occurred within the choice epoch. An analysis of perseverative and regressive errors during reversal blocks revealed the mPFC is critically and differentially involved during the delay and choice epochs of our task. Specifically, this error analysis revealed that when the mPFC was disrupted during delay epochs, animals tended to maintain the previous strategy longer (i.e. showed more perseverative errors), while mPFC disruption during choice epochs tended to cause a failure to retrieve or implement the newly learned strategy (i.e. showed more regressive errors). Lastly, our analysis of VTE behavior showed that disrupting the mPFC during the SSRL task does not impair the average rate of VTE behavior. Instead, it altered the dynamic fluctuation in VTE rates around changes in task demands. Together, these findings reveal differential mPFC involvement in the three working memory processes, and a role for the mPFC in coordinating deliberative behaviors during flexible memory-guided decision-making.

4.4.1 mPFC is involved in working memory encoding, maintenance, and retrieval during flexible decision-making

Past research on the mPFC suggests it has an active role in the encoding, maintenance, and retrieval of task relevant features in working memory (Luk & Wallis, 2009; S. T. Yang et al., 2014). Evidence for the mPFC's role in working memory encoding comes from studies that report mPFC neurons which selectively fire in response to reward outcomes and stimulus presentations (Horst & Laubach, 2012; Warden & Miller, 2010; S. T. Yang et al., 2014; Y. Yang & Mailman, 2018). In support of the notion that the mPFC may shift roles amongst distinct working memory processes,

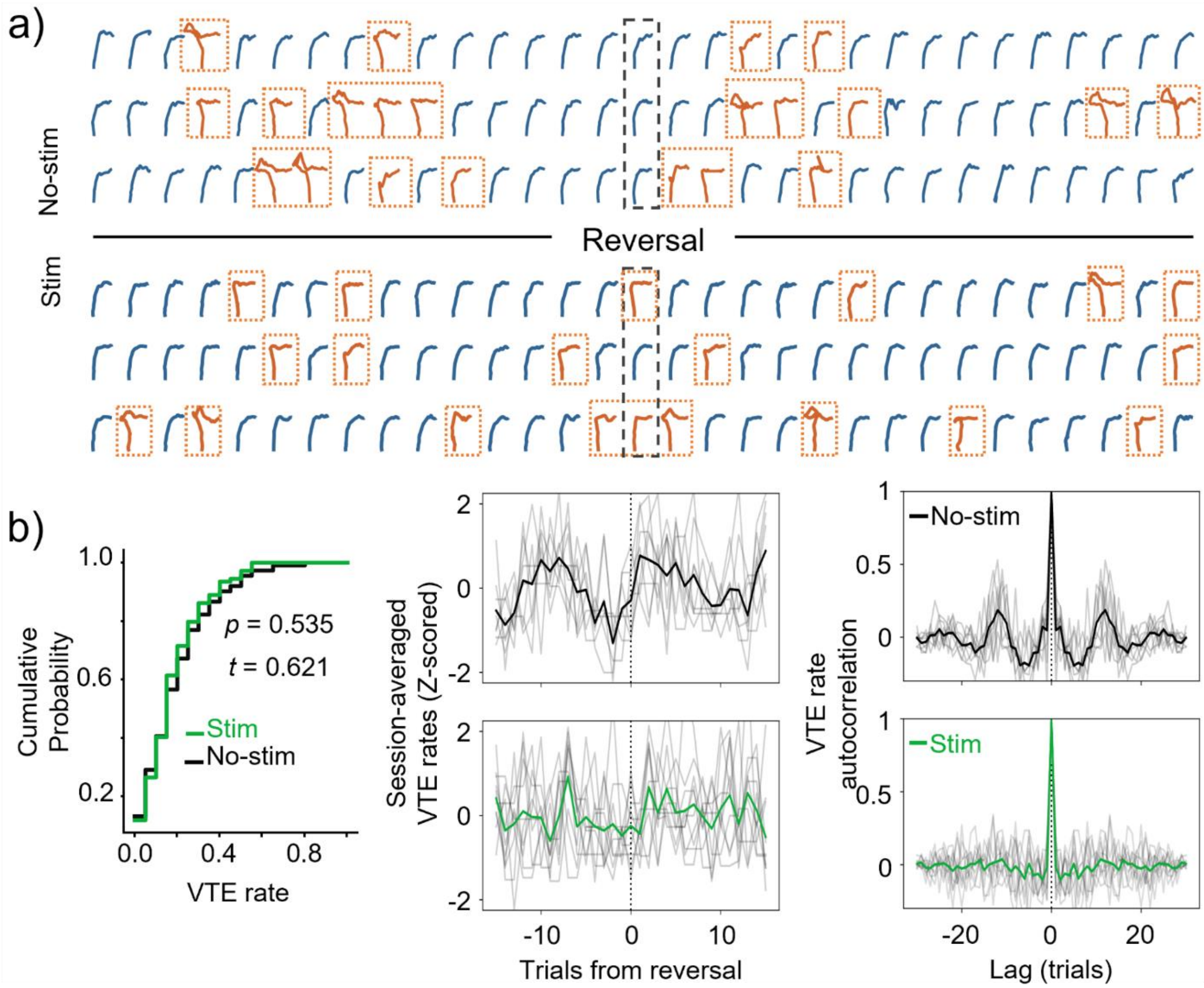

Figure 4.6 Coordination of vicarious trial and error (VTE) behavior with task demand is sensitive to mPFC disruption. The top of panel **(A)** shows sequences of choice trajectories fro m 15, non-stimulated trials before (left of dashed box) and after reversals (right of dashed box) for three different reversal blocks. Choice trajectories showing VTE are colored orange and surrounded by a dotted box. Choice trajectories showing VTE for stimulation trails are shown below the black line. All trajectories are oriented so they start at the bottom and end at the right, and thus do not reflect their actual starting and ending points. In **(B)**, the left panel shows cumulative probability density functions for VTE rates (calculated on reversal blocks) during stimulated sessions in green and non-stimulated sessions in black. There is no significant difference in the distribution of block-averaged VTE rates between stim and no-stim conditions (p = 0.535, t = 0.621, two-tailed, two-sample T-test). The middle panel shows session-averaged VTE rates aligned to reversals, with non-stimulated sessions on top (black) and stimulated sessions on bottom (green). Thin, gray lines are individual session averages (across reversal blocks) and bold, colored lines are the averages across individual sessions. On the right are autocorrelation functions for each respective VTE rate time-series from the middle panel. Note the clear periodicity in VTE rates as a function of trials from reversal when no stim is applied, and general adherence to mean rates when the mPFC is stimulated.

Lundqvist et al. (2016) found increases in the rate of mPFC gamma bursts during stimulus presentations while finding both a decrease in gamma bursts and increase in beta bursts during delay periods. This shift from gamma bursts to beta bursts may be reflective of the mPFC switching between the processes of working memory encoding during reward periods to working memory maintenance during delay periods (Jayachandran et al., 2022; Lundqvist et al., 2016). Also during delay periods, neuronal responses that are selective to spatial locations and choice outcomes can be seen with persistent increased activity (Bolkan et al., 2017; Liu et al., 2014), which further suggests the mPFC maintains information in working memory across delays periods. A role for the mPFC in working memory retrieval comes from the idea that choice points of tasks are times when information must be retrieved and shared across structures to guide action selection (i.e., choice): studies report increased firing rates in mPFC single-units surrounding choice points and LFP recordings show mPFC theta power peaks during choice points (Luk & Wallis, 2009; S. T. Yang et al., 2014; Y. Yang & Mailman, 2018). Additionally, HPC-mPFC oscillatory coherence in the theta band peaks during choice points (Benchenane et al., 2010; Griffin, 2021; Jones & Wilson, 2005; Tamura et al., 2017), suggesting that choice points are critical times when the mPFC communicates with the HPC, possibly to retrieve and share memories which guide action-selection. In further support of this idea, studies (Guise & Shapiro, 2017; Schmidt et al., 2019) report mPFC pharmacological inactivation impairs spatial working memory performance through a disruption of pattern separation by hippocampal prospective codes. This result indicates that the mPFC is necessary to help inform hippocampal representations about task-relevant information which guides prospective decisions. Altogether, these studies indicate that during decision-making, the mPFC encodes task relevant features into working memory, and maintains them over delay periods to help the hippocampus engage in context appropriate memory retrieval during deliberation.

In the current study, we reasoned that the three working memory processes (encoding, maintenance, retrieval) would be differentially distributed throughout epochs of a trial. The delay epoch of the SSRL task is the time when task-relevant features need to be maintained in working memory so they can be used to guide decisions in future trials. Disrupting the mPFC during the delay epoch resulted in the least overall performance impairment and did not impair performance during discrimination learning. The lack of an effect from delay epoch disruption during ID blocks suggests the mPFC is not required for maintaining reward location in working memory when there are no conflicting recent memories (i.e. before reversals). We then conducted an error analysis during reversal blocks and revealed a slight increase in the overall number of perseverative errors that was only significant when the mPFC was disrupted during the delay epoch. This selective increase in perseverative errors implies it was harder for animals to maintain new information about reward contingencies in working memory during delay epoch disruption, subsequently leading to difficulty learning the new strategy. Interestingly, delay epoch disruption was the only condition which switched the sign of the error index value between ID and reversal blocks, indicating a stimulation-driven redistribution from regressive to perseverative errors after the ID. This pattern was significantly different from the choice epoch error index value. These findings provide direct behavioral evidence that the mPFC is involved in an active maintenance of task-representations (working memory maintenance) only once flexible decision-making is required.

The choice epoch of our SSRL task may represent the time when task-relevant information must be retrieved across the hippocampal-mPFC memory system to aid deliberation and choice selection. Selectively disrupting the mPFC during choice epochs consistently impaired performance across more metrics than any other epoch. Also, choice epoch disruption was the only condition which resulted in impairments both to discrimination learning (ID blocks) and to flexible decision-making (reversal blocks). Lastly, our analysis of errors during reversal blocks revealed a striking increase in regressive errors but not perseverative errors during mPFC choice epoch disruption. This selective increase in regressive errors implies animals had a harder time implementing the new strategy, suggesting a failure to retrieve memories related to the new strategy during mPFC choice epoch disruption. Conversely, the lack of an effect on perseverative errors while disrupting during the choice epoch suggests animals had no problem encoding the new strategy.

The return epoch of the SSRL task began at reward delivery and included reward consumption. Therefore, it included the time when outcome information was first available for encoding. Since return epoch disruption did not impair ID performance, the hippocampus was likely able to encode the initial goal location without the mPFC. This is consistent with studies showing cells in the intermediate hippocampus encode goal locations (Aoki et al., 2019; Pfeiffer, 2022) and place cells in dorsal hippocampus increase firing rates during paths towards goals (Jarzebowski et al., 2022; Tryon et al., 2017). However, our data shows that return epoch disruption did significantly impaired the total number of reversals per session. The lack of an effect from disruption during ID blocks when compared to the decreased number of reversals suggests that animals were impaired in their ability to encode new reward information during reversal blocks. This may have reduced the ability to override previously used strategies during return epoch disruption.

Interestingly, for almost all other behavioral metrics, impairments caused by return epoch disruption tended to show an effect intermediate of those seen from delay and choice epoch disruption. This may be in part because, on the SSRL task, impairments caused by a failure to encode information into working memory could, in theory, manifest in the same way as a failure to retrieve or maintain information in working memory. This is because there would be no memory to maintain or retrieve if it was never encoded in the first place. Another possibility is that our epoch designations may contain some overlap between working memory processes. For example, the return epoch of our task includes a traversal from the reward location to the start platform, which could be a time when information needs to be maintained in working memory. However, if the latter was the case we would expect to see performance deficits resulting from return epoch disruption to be more similar to those caused by delay epoch disruption. Our data however show the opposite: return epoch disruption effects are closer in magnitude to those caused by choice epoch disruption. Thus, it does not appear that there is meaningful overlap of working memory processes due to our epoch designations.

The delay epoch of the SSRL task is the time when information must be maintained in working memory. Much like our previous study (Kidder et al., 2021), we did not see any effects of delay epoch mPFC disruption during ID blocks. We reasoned that the mPFC may not have been involved during the delay epoch of the SDA task because it did not require animals to maintain outcome information in working memory. Since the present study's SSRL task did require animals to know which location was rewarded, the lack of an effect from mPFC disruption during delay epochs, specifically in ID blocks, reveals the mPFC is not necessary to maintain spatial or outcome information in working memory when there are no recent conflicting memories.

Our results suggest that the mPFC performs working memory retrieval and/or action selection during basic discrimination learning, which is a time when only one memory representation (goal location) of the task is present. During reversal blocks, a time when multiple recent memories must be considered, mPFC disruption in each epoch impaired performance in different ways. The selective increase in regressive errors during choice epoch disruption suggests the mPFC is important for retrieving recent memories to make decisions. The selective increase in perseverative errors resulting from delay epoch disruption suggests the mPFC has a role in maintaining current strategies during delays only when task representations require updating. Lastly, impairment to the number of reversals completed as a result of return epoch disruption suggests the mPFC supports the encoding of reward information when recent memory conflicts with current task demands. Therefore, these findings reveal that the mPFC becomes critically involved in all working memory processes when animals face decisions requiring the flexible use of memory to discriminate between or update conflicting memories.

4.4.2 The mPFC engages deliberative behavior according to task demands

Vicarious-trial-and-error (VTE) is a behavior in which animals vacillate between path options, presumably in an act of deliberation (Redish, 2016; Schmidt et al., 2013). In Kidder et al. (2021) we found that mPFC disruption, regardless of epoch, decreased the occurrence of VTEs. Furthermore, the decrease in the occurrence of VTEs significantly correlated with the decrease in performance (choice accuracy) caused by mPFC disruption. This demonstrated that 1) the mPFC modulates VTE behavior, and 2) VTEs are correlated with choice behavior. Numerous other studies also demonstrate VTE behavior is dependent on the mPFC and its functional interaction with dorsal hippocampus (McLaughlin et al., 2021; Papale et al., 2016; Schmidt et al., 2019).

One model suggests VTEs may act as a compensatory mechanism to aid deliberation in the face of uncertainty (Amemiya & Redish, 2016; Papale et al., 2012). In our task, uncertainty increases when recent memory conflicts with current task demands (i.e. at reversals), which is when we saw transient increases in VTE rates. Interestingly, in contrast with our previous findings (Kidder et al., 2021), the overall prevalence of VTEs did not change between baseline and stimulation conditions. This could be due to the fact that, during the SDA task, success is only determined by which reward location was immediately visited, not whether it was rewarded, and where the reward contingency never changes, resulting in constant certainty levels. In other words, by introducing periods of uncertainty into the task, we see that VTE rates do not go down when we disrupt the mPFC. Instead, disrupting the mPFC during the SSRL task decouples VTE rates from periods of task-related uncertainty. Thus, in addition to providing further support that the mPFC is causally involved in regulating VTE rates, our current results suggest that a critical role of the mPFC is adjusting when deliberative behaviors occur with respect to changes in uncertainty. Importantly, this explanation does not require undisturbed mPFC involvement for setting the overall VTE rate, which could be a function of uncertainty calculations done elsewhere in the brain (Fiorillo et al., 2003).

Recent results from Mclaughlin and Redish (2023) are also consistent with the interpretation that the mPFC is important for regulating VTE rates with respect to changing uncertainty. The authors show that optogenetically disrupting the mPFC during choices in a spatial delayed discounting task decreases VTE rates, and that these decreases were specific to trials that were close to when rats chose to stop adjusting delay durations for a larger reward. Control rats generally showed higher VTE rates during this period than stimulated rats. Taken together with their other results, the authors use these decreased VTE rates during mPFC disruption to suggest that the mPFC is important for flexibly updating decision-making from deliberative to procedural strategies, which could be another way of describing the decrease in VTE rates that we observed as animals progressed through a block.

4.5 Conclusion

Altogether, this study provides confirmation of past results which suggest mPFC involvement in working memory processes. This study adds to our knowledge by causally revealing mPFC involvement in each working memory process (encoding, maintenance, retrieval) specifically when faced with decisions involving conflicting recent memories which require the flexible use of memories. Importantly, our data revealed a decrease in the number of reversals completed in session during mPFC return epoch disruption, an increase in regressive errors during mPFC choice epoch disruption and, an increase in perseverative errors resulting from delay epoch disruption. These epoch-selective increases in errors are suggestive of distinct working memory processes (encoding, maintenance, and retrieval) being interrupted as a result of mPFC disruption. Our VTE analysis revealed the mPFC is responsible for modulating this behavior as new information needs to be flexibly incorporated into working memory. Overall, our results fit a model in which the mPFC is required for retrieving context appropriate memories during decision-making but only becomes necessary to encode and maintain task relevant information in working memory during flexible decision-making. In this way, the mPFC becomes more engaged in processing working memory information as task demands change and uncertainty increases.

References

Amemiya, S., & Redish, A. D. (2016). Manipulating Decisiveness in Decision Making: Effects of Clonidine on Hippocampal Search Strategies. *The Journal of Neuroscience*, *36*(3), 814–827. https://doi.org/10.1523/JNEUROSCI.2595-15.2016

Aoki, Y., Igata, H., Ikegaya, Y., & Sasaki, T. (2019). The Integration of Goal-Directed Signals onto Spatial Maps of Hippocampal Place Cells. *Cell Reports*, *27*(5), 1516-1527.e5. https://doi.org/10.1016/j.celrep.2019.04.002

Avigan, P. D., Cammack, K., & Shapiro, M. L. (2020). Flexible spatial learning requires both the dorsal and ventral hippocampus and their functional interactions with the prefrontal cortex. *Hippocampus*, *30*(7), 733–744. https://doi.org/10.1002/hipo.23198

Baddeley, A. (2011). Working Memory: Theories, Models, and Controversies. *Annu. Rev. Psychol. 2012*, *63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baeg, E. H., Kim, Y. B., Huh, K., Mook-Jung, I., Kim, H. T., & Jung, M. W. (2003). Dynamics of Population Code for Working Memory in the Prefrontal Cortex. *Neuron*, *40*(1), 177–188. https://doi.org/10.1016/S0896-6273(03)00597-X

Batuev, A. S., Pirogov, A. A., & Orlov, A. A. (1979). Unit activity of the prefrontal cortex during delayed alternation performance in monkey. *Acta Physiologica Academiae Scientiarum Hungaricae*, *53*(3), 345–353.

Becker, J. T., Olton, D. S., Anderson, C. A., & Breitinger, E. R. P. (1981). Cognitive mapping in rats: The role of the hippocampal and frontal systems in retention and reversal. *Behavioural Brain Research*, *3*(1). https://doi.org/10.1016/0166-4328(81)90025-5

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., & Wiener, S. I. (2010). Coherent theta oscillations and reorganization of spike timing in the hippocampal-prefrontal network upon learning. *Neuron*, *66*(6), 921–936. https://doi.org/10.1016/J.NEURON.2010.05.013

Birrell, J. M., & Brown, V. J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *20*(11), 4320–4324. https://doi.org/10.1523/JNEUROSCI.20-11-04320.2000

Bolkan, S. S., Stujenske, J. M., Parnaudeau, S., Spellman, T. J., Rauffenbart, C., Abbas, A. I., Harris, A. Z., Gordon, J. A., & Kellendonk, C. (2017). Thalamic projections sustain prefrontal activity during working memory maintenance. *Nature Neuroscience*, *20*(7), 987–996. https://doi.org/10.1038/NN.4568

Buzsáki, G., Mckenzie, S., & Davachi, L. (2021). *Annual Review of Psychology Neurophysiology of Remembering*. https://doi.org/10.1146/annurev-psych-021721

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*(6625). https://doi.org/10.1038/386604a0

De Bruin, J. P. C., Feenstra, M. G. P., Broersen, L. M., Van Leeuwen, M., Arens, C., De Vries, S., & Joosten, R. N. J. M. A. (2000). Role of the prefrontal cortex of the rat in learning and decision making: effects of transient inactivation. *Progress in Brain Research*, *126*, 103–113. https://doi.org/10.1016/S0079-6123(00)26010-X

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science (New York, N.Y.)*, *299*(5614), 1898–1902. https://doi.org/10.1126/SCIENCE.1077349

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349. https://doi.org/10.1152/JN.1989.61.2.331

Fuster, J. M., & Alexander, G. E. (1971). Neuron Activity Related to Short-Term Memory. *Science*, *173*(3997), 652–654. https://doi.org/10.1126/science.173.3997.652

Goldman-Rakic, P. S. (1995). Cellular basis of working memory. *Neuron*, *14*(3), 477–485. https://doi.org/10.1016/0896-6273(95)90304-6

Griffin, A. L. (2021). The nucleus reuniens orchestrates prefrontal-hippocampal synchrony during spatial working memory. *Neuroscience & Biobehavioral Reviews*, *128*, 415–420. https://doi.org/10.1016/J.NEUBIOREV.2021.05.033

Guise, K. G., & Shapiro, M. L. (2017). Medial Prefrontal Cortex Reduces Memory Interference by Modifying Hippocampal Encoding. *Neuron*, *94*(1), 183-192.e8. https://doi.org/10.1016/j.neuron.2017.03.011

Hasz, B. M., & Redish, A. D. (2020). Dorsomedial prefrontal cortex and hippocampus represent strategic context even while simultaneously changing representation throughout a task session. *Neurobiology of Learning and Memory*, *171*. https://doi.org/10.1016/J.NLM.2020.107215

Horst, N. K., & Laubach, M. (2012). Working with memory: evidence for a role for the medial prefrontal cortex in performance monitoring during spatial delayed alternation. *Journal of Neurophysiology*, *108*(12), 3276–3288. https://doi.org/10.1152/jn.01192.2011

Ito, H. T., Zhang, S. J., Witter, M. P., Moser, E. I., & Moser, M. B. (2015). A prefrontal-thalamo-hippocampal circuit for goal-directed spatial navigation. *Nature*, *522*(7554), 50–55. https://doi.org/10.1038/NATURE14396

Jarzebowski, P., Hay, Y. A., Grewe, B. F., & Paulsen, O. (2022). Different encoding of reward location in dorsal and intermediate hippocampus. *Current Biology*, *32*(4), 834-841.e5. https://doi.org/10.1016/j.cub.2021.12.024

Jayachandran, M., Viena, T. D., Garcia, A., Vasallo Veliz, A., Leyva, S., Vertes, R. P., & Allen, T. A. (2022). Reuniens transiently synchronizes memory networks at beta frequencies *BioRxiv*. https://doi.org/10.1101/2022.06.21.497087

Jones, M. W., & Wilson, M. A. (2005). Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. *PLoS Biology*, *3*(12), e402. https://doi.org/10.1371/journal.pbio.0030402

Jung, M. W., Qin, Y., McNaughton, B. L., & Barnes, C. A. (1998). Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. *Cerebral Cortex*, *8*(5), 437–450. https://doi.org/10.1093/CERCOR/8.5.437

Kamigaki, T., & Dan, Y. (2017). Delay activity of specific prefrontal interneuron subtypes modulates memory-guided behavior. *Nature Neuroscience*, *20*(6), 854–863. https://doi.org/10.1038/nn.4554

Kidder, K. S., Miles, J. T., Baker, P. M., Hones, V. I., Gire, D. H., & Mizumori, S. J. Y. (2021). A selective role for the mPFC during choice and deliberation, but not spatial memory retention over short delays. *Hippocampus*, *31*(7), 690–700. https://doi.org/10.1002/hipo.23306

Kinoshita, S., Yokoyama, C., Masaki, D., Yamashita, T., Tsuchida, H., Nakatomi, Y., & Fukui, K. (2008). Effects of rat medial prefrontal cortex lesions on olfactory serial reversal and delayed alternation tasks. *Neuroscience Research*, *60*(2), 213–218. https://doi.org/10.1016/J.NEURES.2007.10.012

Kubota, K., & Niki, H. (1971). Prefrontal cortical unit activity and delayed alternation performance in monkeys. *Journal of Neurophysiology*, *34*(3), 337–347. https://doi.org/10.1152/JN.1971.34.3.337

Liu, D., Gu, X., Zhu, J., Zhang, X., Han, Z., Yan, W., Cheng, Q., Hao, J., Fan, H., Hou, R., Chen, Z., Chen, Y., & Li, C. T. (2014). Medial prefrontal activity during delay period contributes to learning of a working memory task. *Science (New York, N.Y.)*, *346*(6208), 458–463. https://doi.org/10.1126/SCIENCE.1256573

Luk, C.-H., & Wallis, J. D. (2009). Dynamic Encoding of Responses and Outcomes by Neurons in Medial Prefrontal Cortex. *Journal of Neuroscience*, *29*(23), 7526–7539. https://doi.org/10.1523/JNEUROSCI.0386-09.2009

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2016). Gamma and Beta Bursts Underlie Working Memory. *Neuron*, *90*(1), 152–164. https://doi.org/10.1016/j.neuron.2016.02.028

McLaughlin, A. E., Diehl, G. W., & Redish, A. D. (2021). Potential roles of the rodent medial prefrontal cortex in conflict resolution between multiple decision-making systems. 249–281. https://doi.org/10.1016/bs.irn.2020.11.009

Mclaughlin, A. E., & Redish, A. D. (2023). Optogenetic disruption of the prelimbic cortex alters long-term decision strategy but not valuation on a spatial delay discounting task. *Neurobiology of Learning and Memory*, *200*, 1074–7427. https://doi.org/10.1016/j.nlm.2023.107734

Papale, A. E., Stott, J. J., Powell, N. J., Regier, P. S., & Redish, A. D. (2012). Interactions between deliberation and delay-discounting in rats. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(3), 513–526. https://doi.org/10.3758/s13415-012-0097-7

Papale, A. E., Zielinski, M. C., Frank, L. M., Jadhav, S. P., & Redish, A. D. (2016). Interplay between Hippocampal Sharp-Wave-Ripple Events and Vicarious Trial and Error Behaviors in Decision Making. *Neuron*, *92*(5), 975–982. https://doi.org/10.1016/j.neuron.2016.10.028

Pfeiffer, B. E. (2022). Spatial Learning Drives Rapid Goal Representation in Hippocampal Ripples without Place Field Accumulation or Goal-Oriented Theta Sequences. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *42*(19), 3975–3988. https://doi.org/10.1523/JNEUROSCI.2479-21.2022

Pratt, W. E., & Mizumori, S. J. Y. (2001). Neurons in rat medial prefrontal cortex show anticipatory rate changes to predictable differential rewards in a spatial memory task. *Behavioural Brain Research*, *123*(2), 165–183. https://doi.org/10.1016/S0166-4328(01)00204-2

Ragozzino, M. E., Detrick, S., & Kesner, R. P. (1999). Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(11), 4585–4594. https://doi.org/10.1523/JNEUROSCI.19-11-04585.1999

Redish, A. D. (2016). *Vicarious trial and error. Nat. Rev Neuro.* https://doi.org/10.1038/nrn.2015.30

Samborska, V., Butler, J. L., Walton, M. E., Behrens, T. E. J., & Akam, T. (2022). Complementary task representations in hippocampus and prefrontal cortex for generalizing the structure of problems. *Nature Neuroscience*, *25*(10), 1314–1326. https://doi.org/10.1038/S41593-022-01149-8

Schmidt, B., Duin, A. A., & Redish, A. D. (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. *Journal of Neurophysiology*, *121*(6), 1981–2000. https://doi.org/10.1152/jn.00793.2018

Schmidt, B., Papale, A., Redish, A. D., & Markus, E. J. (2013). Conflict between place and response navigation strategies: Effects on vicarious trial and error (VTE) behaviors. *Learning & Memory*, *20*(3), 130–138. https://doi.org/10.1101/lm.028753.112

Stout, J. J., & Griffin, A. L. (2020). Representations of On-Going Behavior and Future Actions During a Spatial Working Memory Task by a High Firing-Rate Population of Medial Prefrontal Cortex Neurons. *Frontiers in Behavioral Neuroscience*, *14*, 151. https://doi.org/10.3389/FNBEH.2020.00151/BIBTEX

Tamura, M., Spellman, T. J., Rosen, A. M., Gogos, J. A., & Gordon, J. A. (2017). Hippocampal-prefrontal theta-gamma coupling during performance of a spatial working memory task. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-02108-9

Tryon, V. L., Penner, M. R., Heide, S. W., King, H. O., Larkin, J., & Mizumori, S. J. Y. (2017). Hippocampal neural activity reflects the economy of choices during goal-directed navigation. *Hippocampus*, *27*(7), 743–758. https://doi.org/10.1002/HIPO.22720

Warden, M. R., & Miller, E. K. (2010). Task-dependent changes in short-term memory in the prefrontal cortex. *Journal of Neuroscience*, *30*(47). https://doi.org/10.1523/JNEUROSCI.1569-10.2010

Yang, S. T., Shi, Y., Wang, Q., Peng, J. Y., & Li, B. M. (2014). Neuronal representation of working memory in the medial prefrontal cortex of rats. *Molecular Brain*, *7*(1). https://doi.org/10.1186/S13041-014-0061-2

Yang, Y., & Mailman, R. B. (2018). Strategic neuronal encoding in medial prefrontal cortex of spatial working memory in the T-maze. *Behavioural Brain Research*, *343*, 50–60. https://doi.org/10.1016/j.bbr.2018.01.020

Zylberberg, J., & Strowbridge, B. W. (2017). Mechanisms of Persistent Activity in Cortical Circuits: Possible Neural Substrates for Working Memory. *Annual Review of Neuroscience*, *40*. https://doi.org/10.1146/annurev-neuro-070815-014006

Chapter 5

Flexible decision-making is related to vicarious trial and error and strategy learning during spatial set-shifting

Abstract

A hallmark of behavioral flexibility is the ability to update behavior in response to changes in context. Studies of behavioral flexibility classically employ reversal learning or set shifting tasks that require subjects to change their behaviors in response to reward contingency or rule changes. Flexibility is often measured by the number of contingency or rule switches experienced, how many trials it took to complete a block of a certain type, or the types of errors exhibited within a block, which, at their core, tend to reduce to calculations of choice accuracy over periods of several or many trials. However, in contexts where multiple possible strategies might be employed to solve a task, flexibility could manifest as switching between different strategy types that are not necessarily correct – a behavior that could be obscured by reliance on choice accuracy. What's more, these measures are difficult to adapt in a way that allows trial-by-trial estimates of flexibility. Animal and human subjects have also been shown to exhibit hesitation and decision reversals throughout the course of a decision, a behavior often called vicarious trial and error (VTE) in rodent literature and change of mind in primate literature. To understand how these complementary types of information about decision-making relate to learning and behavioral flexibility on both the single trial scale and as different reward contingencies are learned, we assessed strategy use with a previously developed recency-weighted Bayesian inference algorithm that models the likelihood a subject is using an explicitly defined strategy on a trial-by-trial basis. These estimates helped identify within-block learning phases and were used as the basis for a behavioral flexibility score, defined by changes in strategy likelihoods between trials. Aligning flexibility scores to learning points showed that flexibility increases coincided with learning, peaking just after estimated learning points. Additionally, learning dynamics were related to VTE behavior in that VTEs were closely locked to estimated learning points, and, consequently, behavioral flexibility. We complemented these observations by showing that correct VTEs were also more likely to have higher flexibility scores than incorrect VTEs, but that there were certainly examples of VTE during inflexible behavior. Interestingly, VTEs during inflexible behavior were more likely to be wrong than VTEs during flexible behavior. Overall, we demonstrated the use of multiple measures that could assess relationships between learning, behavioral flexibility, and decision-making behaviors. Further, we used these complementary measures to demonstrate that a particular decision-making behavior, VTE, was likely to be a marker of deliberation at some times, and indecision or uncertainty at others.

5.1. Introduction

Behavioral flexibility describes the ability to change behavior in response to changing external conditions or internal states (Brown & Tait, 2010; Dalley et al., 2004; Hones & Mizumori, 2022; Izquierdo et al., 2017; Ragozzino, 2007; Uddin, 2021). Typical tests of behavioral flexibility involve assessing how well subjects perform tasks that require them to abandon previously learned contingencies or rules and implement new behavioral strategies that confer success. In humans, one famous example is the Wisconsin Card Sorting Test (WCST), which requires subjects to learn which stimulus quality (color, number, or shape) is rewarded, match/sort cards to the appropriate quality, and update their sorting strategies as rewarded qualities switch between features (Grant & Berg, 1948; Miyake et al., 2000; Uddin, 2021). Similar tasks have been adapted to non-human primates (Butter, 1969; Goudar et al., 2023; Mahut, 1971; Moore et al., 2005; Roberts et al., 1988), and rodents (Becker et al., 1981; Izquierdo & Jentsch, 2012; Kolb et al., 1974; Ragozzino et al., 2003), though often with simplifications.

A common type of behavioral flexibility test in rodents is the reversal learning task, where an initial reward contingency is suddenly changed, often such that an opposite response or stimulus selection confers reward. For example, rats may have to choose between two potential reward locations, where rewards may initially be located on the right arm of a T-maze, followed by a contingency reversal that gives rewards on the left arm. The reversal requires recognition that the prior location is no longer rewarded, and updating behavior to check whether the other location will yield reward. A related category of behavioral flexibility test is the set shifting paradigm. Set shifting is often characterized by a change in the task rule that dictates reward contingency. Building on the previous example, set shifting on a task with two possible reward locations might first require rats to continually return to the same location (say, the right arm) at the beginning of a session, and then alternate between arms every trial later in the session. The rule changes from “continually return to the same location” to “alternate between locations”. In this case the prior response can be rewarded, and a simple reversal of the behavioral response wouldn’t lead to success.

Flexibility on both reversal learning and set shifting tasks is often measured by 1) number of trials to a certain performance criterion; 2) number of errors due to use of a strategy that’s no longer rewarded, called perseverative errors; or 3) overall choice accuracy within blocks of trials, sometimes broken down by proximity to switches. Additionally, researchers have demonstrated that rodents (Muenzinger, 1938; Muenzinger & Gentry, 1931; Tolman, 1926), non-human primates (Medin et al., 1970), and humans (Santos-Pata & Verschure, 2018; Voss & Cohen, 2017) will sometimes appear to change the course of a decision as its being made, a behavior typically called vicarious trial and error (VTE) but sometimes known as change of mind in primate literature (Kaufman et al., 2015; Resulaj et al., 2009). Often interpreted as a marker of deliberative behavior (Redish, 2016), VTE could serve as another candidate for assessing behavioral flexibility.

Multiple studies have shown that, during tests of flexible decision-making using neuroeconomic paradigms, VTEs tend to increase when rats used their decisions to adjust a delay period in a delay discounting task (Bett et al., 2015; Kreher et al., 2019; McLaughlin & Redish, 2023; Papale et al., 2012, 2016). Similarly, in a “restaurant row” task with multiple reward options and delay durations, VTEs increased when rats were faced with choices near their delay threshold for a particular reward option (Schmidt et al., 2019; Steiner & Redish, 2014). In both tasks, periods with more difficult decisions were more likely to show VTE, but manipulating different parts of the brain changed the relationship between VTE and decision-making in different ways. When the perirhinal cortex – a cortical intermediary between the hippocampus and prefrontal cortex – was silenced, VTEs increased during the later phase of the delay discounting task, when they were typically decreased as a different strategy was exploited (Kreher et al., 2019). Conversely, lesioning the hippocampus decreased the number of VTEs during adjustment periods, but the effect was not significant (Bett et al., 2015). Transiently disrupting the medial prefrontal cortex (mPFC) significantly decreased stable exploitation behavior (prolonged adjustments), while also decreasing VTE in late parts of the delay discounting task (McLaughlin & Redish, 2023). Manipulating the medial prefrontal cortex in the restaurant row task decreased the likelihood of VTEs on difficult decisions (Schmidt et al., 2019). Together, these results suggest that VTE and other aspects of flexible decision-making interact, and manipulating different parts of the neural circuits that support these behaviors can lead to complex breakdowns of these interactions and the behaviors themselves.

Different types of brain manipulation also affect decision-making behavior and performance differently on reversal learning and set shifting tasks. Rats with amygdala lesions increased their VTE and error rates over the course of multiple sessions when reward locations reversed on a T-maze, but un-lesioned controls, on the other hand, showed decreased rates of both over time (Kemble & Beckman, 1970). An allocentric, plus-maze version of this task using within-session reversals from East to West reward locations also showed that VTE rates were highly sensitive to changes in reward contingency reversals, and that brief optogenetic disruptions of the medial prefrontal cortex (mPFC) abolished the tight temporal coupling between VTE rate changes and reward contingency changes (Kidder et al., 2023). Interestingly, although VTE timing became unrelated to task demands, overall VTE rates remained the same. Hippocampal lesions prevented the typical increase in VTE rates seen when rewards were moved from their initial location in a multiple Y-maze with several possible reward sites (Bett et al., 2012). These studies show that VTEs typically change with respect to task demands and task learning, but manipulating different parts of the brain can have different effects on when and how often VTE occurs.

In a multiple T-maze task that instead required rats to *either* perform reversals (as above) *or* switch between egocentric place and alternation rules between sessions, systemic administration of an NMDA receptor antagonist decreased choice accuracy and VTE rates early in sessions (Blumenthal et al., 2011). In contrast, rats with dorsal or ventral hippocampi silenced on a single T-maze that required reversals within (instead of between) sessions did not show differences in overall choice accuracy, but the rats with dorsal hippocampi silenced took longer to switch to alternation strategies after switches from place contingencies (Meyer-Mueller et al., 2020). Plus, although VTEs typically occurred as the new contingency was learned and were associated with incorrect choices, silencing the dorsal hippocampi delayed increases in VTE rates until later in the block, when performance for the contingency had reached criterion. When rats had to switch between egocentric and allocentric navigation strategies on a plus-maze, VTEs increased during allocentric navigation, and were more likely to lead to errors in most circumstances (Schmidt et al., 2013). That said, VTEs did happen during periods of high task proficiency, especially on trials where different strategies would have conflicted with one another.

What seems unequivocal (at least when the brain is not manipulated) is that VTE rates increase when behaviors need to be flexibly updated based on changing task demands. Even so, the links between VTE and choice outcome are not always clear or consistent across tasks. There are several hypotheses about what VTE is and why it happens that we can consider when trying to reconcile these results. Most of the initial studies on, or observations of, VTE showed that VTE tended to happen just before or as rats learned a task, as measured by sharp changes in error rates (Gentry, 1930; Muenzinger, 1938; Muenzinger & Gentry, 1931; Tolman, 1926). Muenzinger (1938) and Gentry (1930) noted that, although they had assumed VTE would be primarily used for comparing sensory stimuli, rats still showed VTE when sensory environments were the same and thus not useful in determining where to go for reward. This suggested that the behavior was not reliant on sampling and comparing sensory stimuli but may indicate comparison of prior experiences to the current situation. Additionally, rats would continue to VTE throughout their learning and training during difficult tasks, but they would typically stop after they had learned to consistently make simple sensory discriminations, often interpreted as having formed a habit (Gentry, 1930; Muenzinger, 1938; Tolman, 1948).

In the context of more recent experiments, the repeated observations that VTE tends to occur early in a block of trials with a new reward contingency or rule, and decrease later into blocks (Blumenthal et al., 2011; Kidder et al., 2023; Meyer-Mueller et al., 2020), is in accordance with the early observations that VTE is linked to learning. Inconsistency in how VTE relates to choice outcome (George et al., 2023; Kidder et al., 2021; Meyer-Mueller et al., 2020; Schmidt et al., 2013), however, suggests that what VTE represents or is used for may not have a unitary explanation (Goss & Wischner, 1956). This possibility was reported in Gentry (1930), who showed that some rats seemed to VTE consistently while never learning proficiently, while others performed exceptionally well, but did not exhibit the typical decline in VTE rates. Her characterization was that VTE consistently associated with poor performance could indicate never having truly learned the task, while VTE during high performance marks continued deliberation. Tolman similarly claimed that VTE during difficult sensory discriminations may persist because comparison and indecision persist, while its increase during initial learning on easy sensory discriminations is because rats are concurrently learning which sensory stimuli (visual/auditory) are associated with reward, as well the discriminative reward contingency (black vs white/toward tone vs away from tone) itself (Tolman, 1948).

Though the evidence discussed so far supports the general claim that VTE is associated with behavioral flexibility, most objective measures of behavioral flexibility can only be evaluated over multi-trial timescales and are reliant on measuring changes in error rates. While there is evidence suggesting the VTE and associated behaviors are affected over these longer timescales (George et al., 2023; McLaughlin & Redish, 2023; Papale et al., 2016), we wanted to measure the association between behavioral flexibility, learning, VTE, and choice outcomes more directly, on a trial-by-trial basis, using behavioral measures that could be calculated independently of one another. To do so, we implemented a spatial set shifting task that required rats to repeatedly switch between learned strategies in which they continually returned to the same location (followed a place rule) or alternated between two locations on every trial (followed an alternation rule). We utilized a previously developed recency-weighted Bayesian inference strategy (Maggi et al., 2022) that compares choice history to explicitly modeled behavioral strategies and updates every trial to compute the likelihood that a given strategy was being used. Using the relative likelihoods in comparison with target strategies, we identified putative learning points for each contingency block, and, using changes in likelihoods across trials, we computed a behavioral flexibility score to determine periods of high flexibility that might not be obvious by proximity to learning points or task structure (e.g., block types/switches) alone.

Our results suggest a typical tight coupling between learning, behavioral flexibility, VTE, and choice outcomes. Increases in choice accuracy, VTE rates, and flexibility scores were all aligned to learning points. Furthermore, VTEs were more likely to lead to correct choices in almost all sessions. Interestingly, VTEs on incorrect choices had distributions of flexibility scores shifted significantly lower than VTEs that led to a correct choice. Complementing this result, VTE trials with low flexibility were more likely to lead to incorrect choices. Taken together, these results support the idea that there are multiple types of VTE, and while the majority of VTEs in the spatial set shifting task appear to be deliberative, some are more indicative of indecision or uncertainty. Overall, these findings help reconcile prior inconsistencies in the VTE literature and provide a new series of corroborating measurements for evaluating the relationships between learning, deliberation, and behavioral flexibility.

5.2. Results

Following several other set shifting or reversal learning paradigms, we utilized a spatial set shifting task (**Figure 5.1A**) that required rats to either continually return to the same location (a place rule) or alternate between locations on successive trials (an alternation rule). This design is similar to Meyer-Mueller et al., (2020), except we use a plus-maze instead of T-maze, with randomly chosen start arms to ensure that only strictly allocentric, place strategies (as opposed to egocentric, body-turn strategies) will be successful. Because prior results show differences in VTE rates for egocentric compared to allocentric navigation (Schmidt et al., 2013), but no consistent changes during switches between different egocentric strategies (Meyer-Mueller et al., 2020), we analyzed whether there were performance differences in allocentric place compared to allocentric alternation strategies. Choice accuracy distributions for all alternation blocks were significantly right-shifted (**Figure 5.1B,** top right; 2-sample, 2-tailed, T-test; t = 2.66; $\overline{alt}$ = 72%, $\overline{place}$ = 69%; p = 0.009; d = 0.42; overbars represent the sample mean). Block duration distributions were non-normally distributed for alternation trials, which showed heavily concentrated cumulative probability bunched around the lower block duration limit (15 trials). This distribution was left-shifted compared to the distribution of place block durations (**Figure 5.1B**, top right; 2-sample, 2-tailed Wilcoxon rank-sum test; Z = -2.94; alt_m = 23.3, $place_m$ = 25.8, where subscript m denotes median; p = 0.003; d = -0.3). Within session comparison of the adjacent pairs of place and alternation blocks (**Figure 5.**1B, bottom right) yields the same result for average choice accuracy difference (t = 2.71; $\overline{\Delta acc}$ = 3.02%; p = 0.008; d = 0.30) and difference in block duration (t = 2.71; $\overline{\Delta dur}$ = -2.78; p = 0.03; d = -0.24). While these differences are consistent, note that they are not large (~3% performance difference and ~3 trial duration difference; Cohen's d values below 0.5).

We implemented a previously developed algorithm that uses recency weighted Bayesian inference to identify changes in strategy learning (Maggi et al., 2022). This algorithm, which assesses strategy use likelihood by comparing decision history (not choice outcome) to explicitly modeled

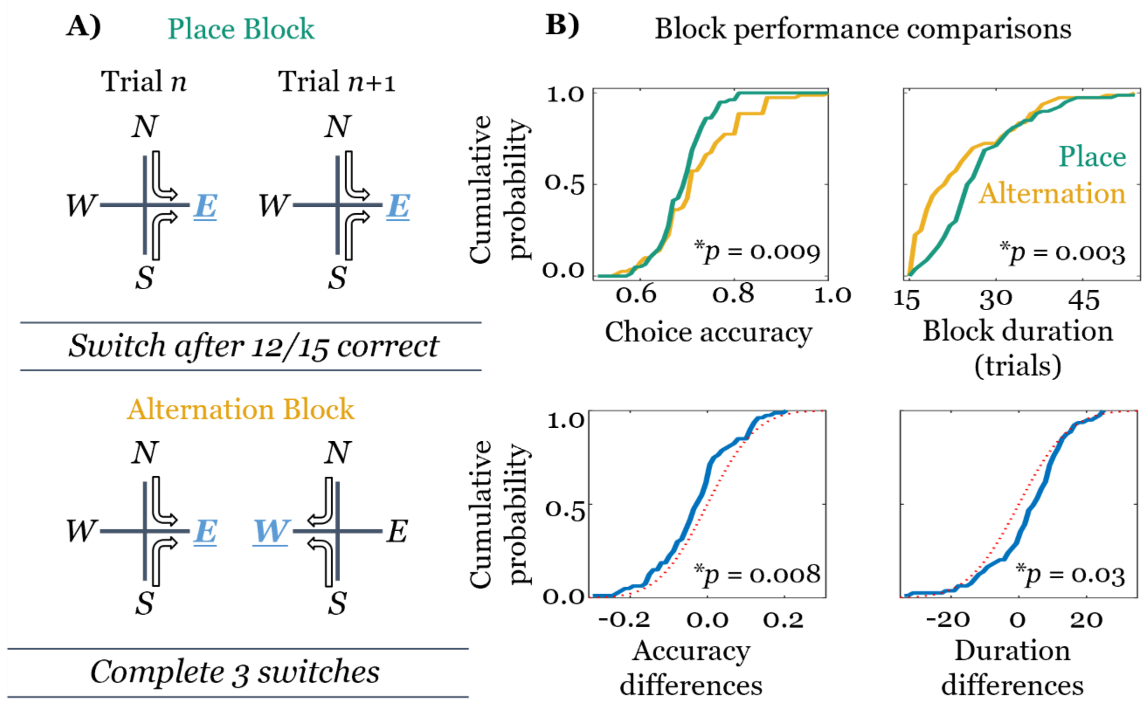

Figure 5.1 *Performance differs according to task rule.* **A)** Diagram of the task rules. Place blocks (top) reward continual visits to a particular (E or W) arm, and alternation blocks (bottom) reward alternation between E and W blocks on successive trials. Sessions consist of 3 block switches, with switches occurring if 12 of the previous 15 choices were consistent with the target rule. **B)** Shows that multiple measures of performance are better for alternation blocks compared to place blocks. Unpaired cumulative distributions for choice accuracy within a block are leftward shifted for place blocks (top left, green line), while alternation block cumulative distributions for block duration are leftward shifted for alternation blocks (top right, gold line). p-values are calculated using unpaired, two-sample, two-sided T-tests. Within-session, paired comparisons suggest the same conclusion (bottom). Differences between choice accuracy in adjacent place and alternation blocks (bottom left solid blue line, place minus alternation) are shifted to the left of zero, while differences in block duration are shifted to the right of zero (bottom right solid blue line, place minus alternation). p-values are calculated using signed rank tests. Red dashed lines are zero-mean/median, standard deviation-matched normal cumulative distributions for comparison.

strategies, allowed us to identify learning by finding when the rat's most likely strategy matched the target task rule (see **Figure 5.2A** for an example of the estimation output). Since the learning point is identified without explicit reference to choice outcome, seeing increases in the likelihood of a correct choice with respect to the putative learning point would corroborate that it had been correctly identified. As expected, average choice accuracy aligned to learning points shows a striking increase just prior to the learning point, remaining elevated for several trials after. As shown in **Figure 5.2B**, average choice accuracy (dashed horizontal line) for the 15 trials up to the learning point (dashed vertical line) is 63.5% and the lower bound of an estimated 95% confidence interval exceeds that value starting 1 trial before the learning point, peaks 1 trial after the learning point, and remains above the pre-learning point average for 9 trials after the learning point (data shown for n = 40, sessions, where each point within 15 trials on either side of the learning point is the average choice accuracy across 4 blocks; same result using n = 13 subjects with averages across 4 to 24 blocks).

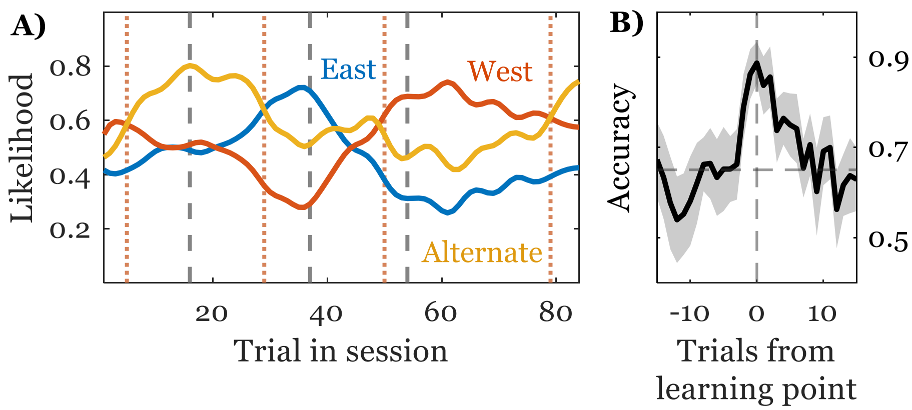

Figure 5.2 *Identifying learning points from strategy likelihoods*. **A)** We modeled three strategies – go east (blue), go west (orange) and alternate (yellow). Learning points are indicated with vertical, orange, dotted lines, and block switches are indicated by vertical, grey, dashed lines. The average choice accuracy aligned to learning points is shown in **B)**. A shaded 95% confidence interval surrounds the average. The horizontal dashed line indicates the pre learning point average, the vertical dashed line indicates the learning point.

As mentioned, prior reports show VTE rate differences for different types of strategy (Schmidt et al., 2013). Both types of strategy had right-skewed, overlapping probability distributions of VTE rates (**Figure 5.3A,** top right; 2-sample, 2-tailed Wilcoxon rank-sum test; Z = 0.36; alt_m = 15%, $place_m$ = 14%; p = 0.72; d = 0.03). Other studies suggest that the relationship between VTE and choice outcome is, if nothing else, task dependent, so we calculated the within session difference between the number of VTEs that led to correct and incorrect choices. For this task, VTEs were far more likely on correct choices. In fact, there were only 4 sessions (10%) where VTEs led to errors more often than correct choices, and out of those 4, VTEs were far more likely to precede errors in only 1 (**Figure 5.3B;** t = 5.04; $\overline{\Delta VTE}$ = 8.8; p < 0.001; d = 0.80, red lines denote sessions with more VTEs leading to errors).

The current and historical literature don't seem to have come to a consensus on how VTEs should unfold throughout the course of learning. Some report that VTE in navigation or location-based tasks decrease over time as learning occurs (Jackson, 1943; Kemble & Beckman, 1970) in a task dependent manner (Goss & Wischner, 1956), but in other tasks VTE has been shown to stay elevated throughout, supposedly depending on the task difficulty (Gentry, 1930; Tolman, 1948). Our task ensures that the current contingency has been learned at the end of a block but is unknown at the beginning of a block. Thus, we asked if there were differences in the number of VTEs in the first 10 trials of a block compared to the last 10. We reasoned that if there were more VTEs during the well learned period, VTEs on our task are likely to be carried out after a contingency has been learned. This would be in line with the hypothesis that VTEs are for deliberation because deliberation requires that subjects understand how different outcomes are related to different choices. It could also be, however, that VTE was more likely early in a block, either because rats have automated their behavior toward the end of a block, or because VTEs were more likely a sign of indecision than deliberation in this task. If there aren't differences, it could be that there were changes within the block that drove changes in VTE (e.g., deliberative behavior prior to or aligned to learning). We saw that there were no differences in the number of

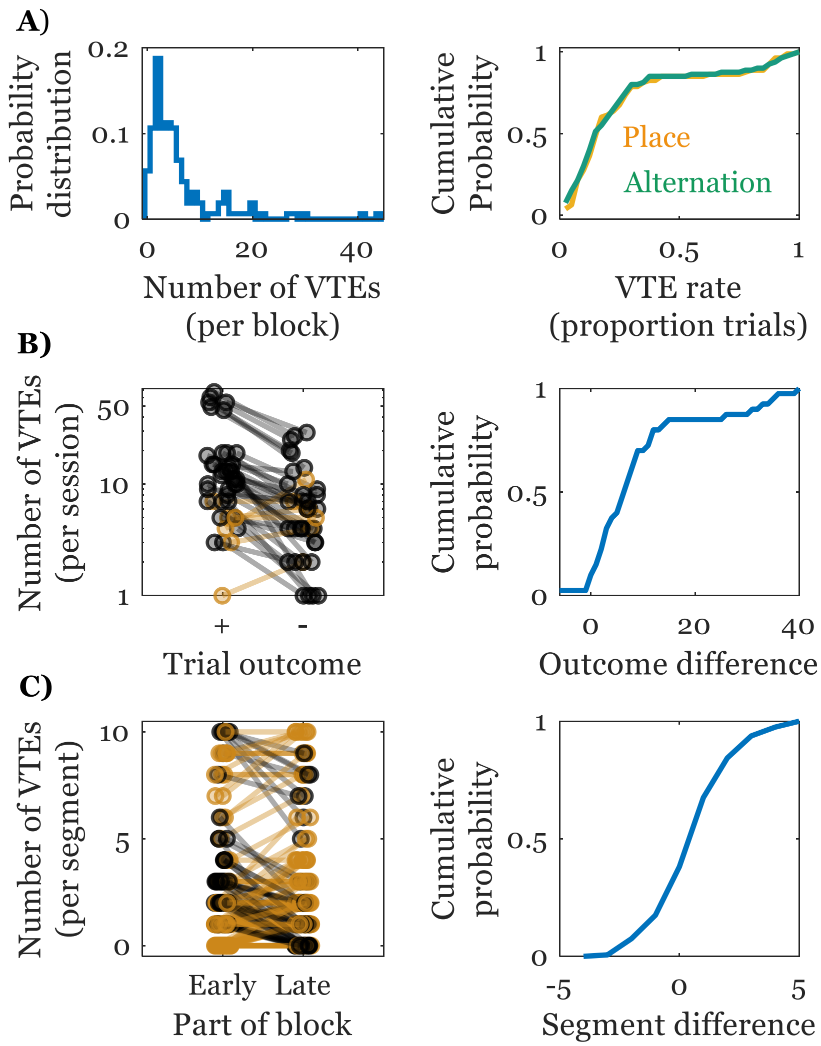

Figure 5.3 *Summary of VTEs by session and block.* **A)** The left histogram shows how the number of VTES are distributed per block, while the right shows proportions of VTEs per block. The right column in **A)** shows that there are no differences in how proportions of VTEs are distributed per block for either alternation (gold) or place (green) blocks. Looking at within session differences in VTEs that lead to correct vs. incorrect choices, **B)** shows that VTEs are far more likely to lead to a correct choice. Raw numbers are shown on the left (with log y-axis on left, not log x on right), and the distribution of within session differences are shown on the right. Although VTEs are more likely to lead to correct choices, **C)** shows that there are no differences in the probabilities of VTEs occurring early compared to late in blocks.

VTE early vs late in the block (**Figure 5.3C;** t = -0.43; $\overline{\Delta VTE}$ = -0.05; p = 0.67; d = -0.03, red lines denote blocks with more VTEs later in the block).

Because VTE has been suggested to track task learning and has been proposed as a behavioral marker for deliberation, we asked whether changes in VTE rates align to learning points. Indeed, **Figure 5.4B** shows that there are significantly elevated VTE rates from 1 trial before to 1 trial after the learning point (see **Methods** for estimation of VTE rates and statistical testing paradigm), further validating the association between learning and VTE without appealing to choice accuracy. These elevations contrast with VTE rates aligned to block switches, which hover around the average for almost the entire window (**Figure 5.4A**).

Changes in strategy likelihoods suggest updates in choice behavior, thus allowing us to define a behavioral flexibility measure based on trial-by-trial strategy likelihood changes (**Figure 5.5A**). Importantly, this method allowed us to measure behavioral flexibility independently of choice outcome, which could mask instances where subjects did, in-fact, switch between strategies, but neither strategy matched the rule they were meant to follow. Further, it allowed us to assess flexibility trial-by-trial instead of over periods of trials. Much like choice accuracy and VTE rate

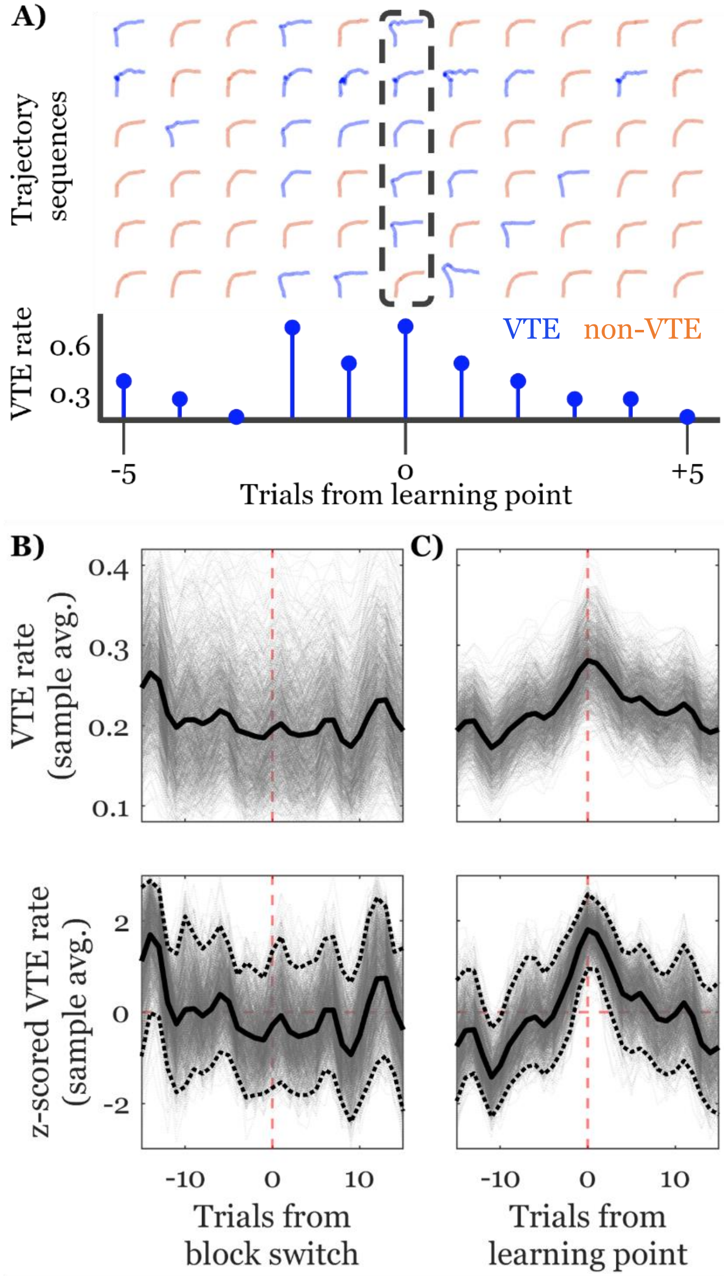

Figure 5.4 *VTE rate changes align with learning points*. Each row in **A)** shows trajectories from sequences of trials aligned to estimated learning points (dashed rectangle/trial 0 on the bottom axis) for a given block. Blue trajectories have been identified as VTEs and orange trajectories have been identified as non-VTEs. The stem plot shows VTE rates for this sample of six sequences. Note the general increased rate for trials near the learning point. The top panels in **B)** and **C)** show changes in VTE rate with respect to block switches on the left and estimated learning points on the right (dashed, vertical, red lines). For all plots, bold, black lines are the averages of 1000 iterations of a hierarchical bootstrap to estimate VTE rate timeseries from binary vectors; light grey lines are the individual iteration results. Top panels show raw rates and bottom plots show z-scored rates. Dashed, horizontal, red lines in the lower plots show the mean for comparison. Dotted black lines show the 2.5 and 97.5 percentiles of the distribution. Dotted black points on the graph above the dashed line at zero indicate that less than 2.5% of the estimates across iterations were below zero for that trial (trials -1 to 2), and vice-versa for points below zero (trial -10).

dynamics, aligning sequences of flexibility scores to the learning points showed that there were significant, stereotyped increases in flexibility starting just before the learning point that end several trials after (**Figure 5.5B**). Note that although this measure is in a way circular (learning points are defined by strategy likelihoods, as are flexibility scores), flexibility scores can and do vary – sometimes dramatically – away from the learning point. Likewise, sometimes flexibility scores were lower during learning points when transitions happened slowly. Thus, this result was expected, but not guaranteed.

The highly correlated associations between VTE rates, flexibility dynamics, and choice accuracy changes provide strong support for the claim that VTE can serve as a marker for deliberation. However, not all VTEs occurred around learning points, some VTEs occurred during periods of

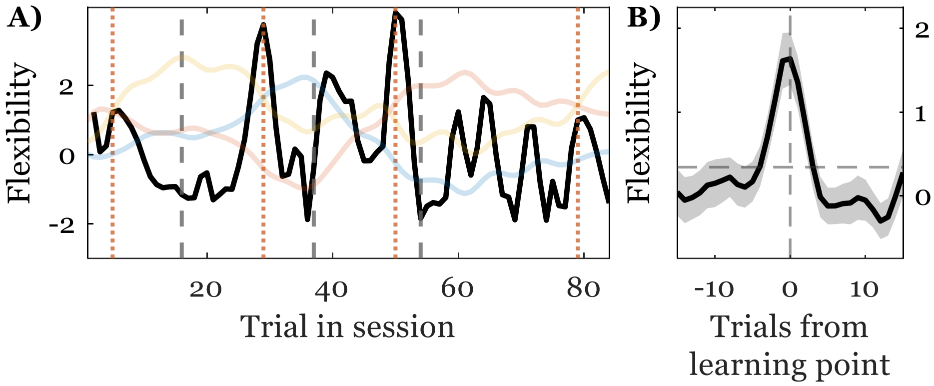

Figure 5.5 *Flexibility score example and relation to learning point.* **A)** Bold, black line indicates flexibility scores for one example session (same as **Figure 5.2A**, shown in background) Learning points are indicated with vertical, orange, dotted lines, and block switches are indicated by vertical, grey, dashed lines. The average flexibility score aligned to learning points is shown in **B)**. A shaded 95% confidence interval surrounds the average. The horizontal dashed line indicates the pre learning point average, the vertical dashed line indicates the

low flexibility, and some VTEs led to errors. Thus, we asked whether there may have been another, non-deliberative type of VTE. We know that VTEs near learning points are happening as choice accuracy and flexibility are high, but to see if opposing relationships existed as well, we asked if incorrect VTEs were associated with lower flexibility scores, and if VTEs that happened in inflexible periods were more likely to be incorrect. Indeed, incorrect VTEs had significantly leftward shifted flexibility scores (**Figure 5.6A**; t = 9.40; $\overline{flex}$ = 0.60, $\overline{inflex}$ = -0.32; $p < 0.0001$; $d = 0.57$). We defined a set of criteria that determined whether a VTE occurred during a flexible or inflexible period. First, any trial within 2 trials of a learning point was considered flexible (regardless of flexibility score). Second, a trial had to be more than 3 trials prior to the end of a block (unless it was within 2 trials from the learning point). Third, any trial with a flexibility score in the top 60th percentile was considered flexible (unless it was within 3 trials from a block switch). Similarly, inflexible periods could not be within 2 trials of the learning point (regardless of flexibility score) and had to have flexibility scores in the bottom 40th percentile (see **Methods** for further descriptions). Within session, paired comparisons showed that choice accuracy was significantly higher during flexible, learning related VTEs than VTEs during inflexible periods that didn't border the learning point (**Figure 5.6B**; t = 5.23; $\overline{\Delta acc}$ = 23.5%; $p < 0.0001$; $d = 1.01$). Together, these data suggest that VTEs can happen under at least two circumstances – one that suggests a deliberative process, and another which looks more like indecision or uncertainty.

5.3 Discussion

Behavioral flexibility is a complex phenomenon that could manifest in many different ways, but our typical understanding of it primarily focuses on a single metric – choice outcome. VTEs have been observed for nearly a century, but it has been difficult to reconcile the literature that has investigated them. This study sought to supplement our understanding of behavioral flexibility and fill in some of the gaps in the VTE literature by analyzing VTE with respect to other streams of behavioral data that also occurred on a trial-by-trial basis during flexible decision-making and dynamic strategy use. To do so, we estimated strategy likelihoods from rule-based models, which enabled learning point identification, and developed a behavioral flexibility measure based on strategy likelihood estimates. We show that choice accuracy, VTEs, and flexibility scores all

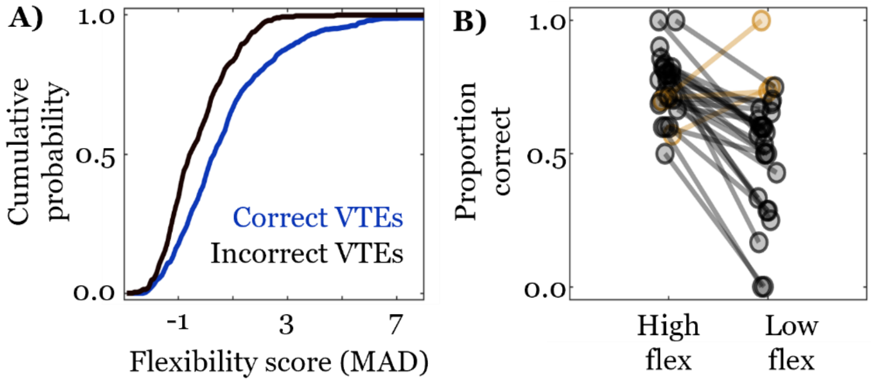

Figure 5.6 *Evidence for multiple VTE types.* **A)** Correct VTEs are more likely to have higher flexibility scores than incorrect VTEs. **B)** VTEs in the top 40th percentile of flexibility scores are more likely to be correct than VTEs in the bottom 40th percentile.

increased surrounding learning points. Further, we show that VTEs were far more likely to be correct than incorrect in this task, and that correct VTEs were more likely to have higher flexibility scores, suggesting a typical role in deliberation. However, we also found VTEs that occurred during periods of low flexibility, which were often wrong, suggesting that VTE may also reflect indecision or uncertainty. Taken together, we show that VTE types varied, but could be separated by using other behavioral measures, such as flexibility scores and choice outcomes.

These results may help explain some of the discrepancies between what other studies have reported about VTE. For example, disrupting the mPFC during a spatial delayed alternation task can cause decreases in VTE rates that correlate with decreases in choice accuracy, which was interpreted as an impairment in deliberation (Kidder et al., 2021). A similar disruption paradigm on a slightly different task that required serial, spatial reversals, however, disrupted VTE timing, but not overall rates (Kidder et al., 2023). Based on what we've reported here, we would interpret this as mPFC disruption halting deliberation, but not inducing indecision, during the spatial delayed alternation task, compared to mPFC disruption causing indecision or uncertainty during the spatial reversal learning task.

Similarly, Schmidt and Redish (2019) showed that long-lasting (i.e., throughout the duration of an experimental session) mPFC disruption decreased deliberative decision-making on difficult choices in the restaurant row task. Interestingly, transient mPFC disruption that only occurred at choice points caused more specific changes in the relationship between VTE and deliberation in the delay discounting task. Namely, it decreased VTE rates late in the task, when rats were typically maintaining their preferred delay, but increased adjustment laps during that same period, when controls would typically maintain instead of adjust (McLaughlin & Redish, 2023). Since adjusting had previously been interpreted as a form of flexible, deliberative decision-making, these authors suggested that the presence of adjustment but lack of VTEs did not strictly indicate a lack of deliberation caused by mPFC disruption. Instead, they proposed that transient mPFC disruption caused difficulty switching from deliberative to procedural behavior. On the other hand, session-long perirhinal inhibition caused similar changes in late-stage adjustment behavior, but *increased* late-stage VTE rates (Kreher et al., 2019). These results were interpreted as causing less stable behavior due to difficulty disambiguating delay-reward associations, which, in turn, was thought to increase deliberation.

We speculate that adjustment during the latter half of the session, when alternation is typical in controls, is not suggestive of deliberation, but is more likely a feature of impaired working memory, which is known to rely on the mPFC (Goldman-Rakic, 1991; Horst & Laubach, 2012; Kidder et al., 2023; Ragozzino et al., 2003; Yang & Mailman, 2018) and mPFC interactions with the hippocampus (Benchenane et al., 2010; Floresco et al., 1997; Hallock et al., 2016; Hyman et al., 2010; Schmidt et al., 2019; Stout et al., 2023; Wang & Cai, 2006; Yoon et al., 2008). The lack of VTEs observed by McLaughlin and Redish could, then, be indicative of a lack of deliberation, while the working memory impairment led to mistimed flexible responding and temporary sub-optimal reward rates. Certainly, these two things could coalesce into inability to shift to a procedural, alternation strategy as suggested by McLaughlin and Redish (2023), and persistently variable flexibility scores late into a session could be evidence for this.

Several mPFC (Kidder et al., 2021, 2023; McLaughlin & Redish, 2023; Schmidt et al., 2019) and hippocampus (Bett et al., 2012, 2015; Blumenthal et al., 2011; Hu & Amsel, 1995; Meyer-Mueller et al., 2020) manipulation studies seem to agree that manipulating these structures is unlikely to increase VTE rates. Curiously, manipulating structures that connect these regions increases VTE (Kemble & Beckman, 1970; Kreher et al., 2019; Stout et al., 2022). One hypothesis for why this may be is that both the hippocampus and mPFC are crucial for specifically enabling deliberative VTE behavior. When one of these structures isn't functioning as it usually does, the deliberation process fails. For example, it may be that sequential hippocampal activity generates possible alternative options (Johnson & Redish, 2007; Kay et al., 2020) while the mPFC evaluates those options prior to choices (Hasz & Redish, 2020; Redish, 2016; Schmidt et al., 2019; Tang et al., 2021; Zielinski et al., 2019). If either the representation of possibilities in the hippocampus or evaluative process in the mPFC fails to occur, VTE may become less likely to happen. In contrast, if both processes proceed as they normally would locally, but are unable to coordinate correctly because of interruptions in connecting circuitry, VTE may be just as likely – if not more likely to occur – but related to indecision or uncertainty instead of deliberation.

As mentioned, behaviors like VTE are present in humans (Santos-Pata & Verschure, 2018; Voss & Cohen, 2017), while inflexible, perseverative, behavior and difficulty with executive control are common measures in clinical diagnoses of neurological disorders (Uddin, 2021). Because any behavior that can be identified and modeled based on decision-making is amenable to the analysis workflow we've utilized, and many decision-making tasks require some active trajectory toward a decision, we hope that the framework described here will be of general interest to behavioral neuroscientists asking both basic and clinical questions.

5.4 Methods

Subjects, apparatus, training protocol, and behavioral task

Food restricted (85% of body weight) Long Evans rats (n = 13, Charles River Laboratories) were trained to perform a spatial set shifting task. Sessions were run on an elevated plus maze (black plexiglass arms, 58 cm long x 5.5 cm wide, elevated 80 cm from floor), with moveable arms and reward feeders controlled by custom LabView 2016 software (National Instruments, Austin, TX, USA) that tracked the rat and automatically raised and lowered arms based on positions recorded by a SONY USB web camera (Sony Corporation, Minato, Tokyo) acquiring frames at approximately 35 Hz. Rats were initially habituated to handlers and the maze for however long it took them to comfortably interact with handlers, forage for pellets on the maze, and habituate to maze movement and noise (roughly several days to a week).

Task training started with a forced choice paradigm, where rats learned the trial structure, consisting of leaving its pseudo-randomly chosen "North" or "South" starting arm, then navigating to an "East" or "West" arm for a 45 mg sucrose pellet reward (TestDiet, Richmond, IN, USA). Initially, rats completed 5 to 7 trials that forced them to either alternate between reward sites or return repeatedly to the same East or West location, regardless of their starting point. This number was increased until rats could do 10 of each reward contingency in under 45 minutes, at which point they were given a free choice version of the same task.

In the free-choice training version, we did not change start arms on error trials until rats made a certain number of correct choices, and initially started with a low choice accuracy criterion for switching between blocks (typically around a 70% success rate in an 8 to 10 trial window). As rats started completing all reward contingencies (alternate, go east, go west), we increased the criterion for success and minimum number of trials per block, and decreased the number of correct trials needed before errors no longer influenced start arm switches. This procedure was tailored to each rat until it could complete 4 reward contingencies (two alternation blocks and two place blocks; one East, one West) in less than 150 trials, with start arms pseudo-randomly chosen for all trials. We also ensured that two alternation blocks did not occur back-to-back.

Testing sessions followed the same trial structure as the final training sessions. As mentioned above, start arms were "pseudo-random". This was done to ensure that between 50 – 60% of trials within 15 trial stretches had start arm switches. Doing so increased the number of switches compared to what you'd expect from random draws, while slightly decreasing the number of long (4 to 10 trial) sequences where the start arm stays the same, and eliminating sequences without a switch that were longer than that. For this dataset, all rats completed three switches, though not all completed the 4th block. A total of 40 sessions were analyzed with all but one rat contributing at least two sessions.

Position tracking and VTE identification

We identified VTEs in much the same way as Kidder et al. (2023). Briefly, we took the videos that tracked coarse body location during the task and used DeepLabCut (DLC) version 2.2 (Mathis et al., 2018; Nath et al., 2019) to identify the rats' heads. We started with the same model trained in Kidder et al. (2023) and retrained a new iteration with additional labeled data from the set shifting experiments. Each training attempt used NVIDIA GEFORCE GTX 1080 GPU with 500,000 iterations. Trajectories with vicarious trial and error (VTE) were detected by projecting the position data into principal component (PC) space and clustering the PC-representations of the trajectories with hierarchical agglomerative clustering. Before projection, all trajectories were aligned and standardized to the same starting and ending positions and interpolated (or linearly subsampled, if necessary) to have the same number of points. Visual inspection of the clustering in PC space naturally formed what looked like two clouds in low dimensional plots, and distance-based dendrograms cut to give two clusters separated trajectories with VTE from non-VTE trajectories, though with some errors.

To mitigate the errors, we reassigned certain trajectories initially not identified as VTE but with x-positions that crossed a certain threshold into the VTE category, and used the combination of a low *z-ln(idphi)* measure (Bett et al., 2012; Blumenthal et al., 2011; McLaughlin & Redish, 2023;

Papale et al., 2012; Schmidt et al., 2013, 2019; Stout et al., 2022) and a failure to cross lower x-position boundary to reassign any VTEs that may have been mistakenly identified. Informal inspections of randomly sampled data subsets after classification suggest that this method is between 80% and 90% accurate, which is in-line with supervised methods and near the threshold for interrater agreement reported in Miles et al., 2021.

Estimating strategy likelihoods, learning points, and flexibility scores

We adopted the procedure from Maggi et al. (2022) to estimate explicitly modeled strategy likelihoods, trial-by-trial, based on choice history and a recency-weighted decay factor to account for the inherent non-stationarity in behavior associated with strategy switching. Minor changes to strategy templates allowed us to model allocentric versions of the strategies instead of the egocentric versions originally used; we simply added a column to the data processing input that parsed "East" and "West" choices instead of "Left" and "Right" (though the algorithm can easily process both types of reference frame if both types of input are given). As in the original report, we used 0.9 as the parameter controlling the strength of the recency weighting. One addition we made was to add several randomly permuted trials to the beginning of sessions prior to processing with the algorithm. This dampened some of the algorithm's initial large swings in likelihood estimates that were due to limited trial history. Further, we smoothed likelihood timeseries with a 5 trial Gaussian window to deal with occasional noisy estimates and increase the reliability of learning point identification. As suggested in the original paper, we identified learning points as the trial when the target strategy became the most likely.

Our rationale for calculating flexibility scores from changes in strategy likelihood is based on the notion that decision-making patterns shifting to become consistent with a different strategy is the sign of flexible behavior in a set shifting task. Thus, for each trial, the flexibility score is the absolute difference in strategy likelihoods from trial t-1 to trial t, summed across strategies. For each session, this value is normalized by median absolute deviation (robust Z-score) because values tend to deviate more dramatically in the positive than negative direction, but the results remain the same when a normal, standard deviation-based Z-score is used.

Flexible periods, used for testing whether there were multiple VTE types were defined, based on three criteria. First, trials on either side of the learning point were automatically considered flexible, regardless of their flexibility score. Trials that were three trials before the end of a block were automatically *not* considered flexible, unless they were within one trial of the learning point. Third, any trials in the top 60% of flexibility scores were considered flexible, and any trials in the bottom 40% of flexibility scores were not considered flexible. Results for this analysis did not change if we changed the ratios of flexible and inflexible VTEs (e.g., based the third criterion on flexibility score instead of percentile), though, of course, this often changed the proportion of data we were able to use.

Statistical quantification

Critical values were set at 0.05, as is tradition. When sample sizes were large, and distributions looked nearly normal, two-tailed T-tests were used. If making within session comparisons, one-sample tests of differences were done, comparing the empirical distribution to what would be expected for a distribution with 0 mean. When distributions were strongly skewed, we performed two-tailed Wilcoxon rank sum tests if the data were not paired, and two-tailed Wilcoxon signed rank tests when they were paired. Learning point aligned choice accuracy and flexibility score data were not subjected to formal statistical testing – instead, 95% confidence intervals were estimated as if they were from a normal distribution (using Z-values). Granted, this method is not particularly precise, but the results were clear and did not seem to need stronger justification.

Learning point aligned VTE “rasters” and a peri-learn point VTE rate averages suggested that VTEs likely aligned to learning as well, but to test this we used hierarchical bootstrapping (Saravanan et al., 2020). This helps safeguard results from bias introduced by uneven data collection between subjects while also acknowledging that multiple measurements from individuals are not independent. It also allowed us to create rate distributions out of binary data. Subjects were sampled randomly with replacement enough times to match the total number of subjects in the dataset (n = 13), then, for each subject, a certain number of blocks was also randomly selected with replacement (8 for learning point and 6 for block switch aligned sequences of VTE data – equivalent to two sessions of data). Average VTE rates were calculated for these samples and smoothed with a 5 trial Gaussian window across trials. Distributions were formed by repeating this procedure 1000 times. Significance was determined by asking which trials had more than 97.5% of their (Z-scored) iterations above 0. Results were tested without smoothing, using larger smoothing windows, with different numbers of iterations, using median absolute deviations instead of standard deviation z-scoring, using shorter and longer sequences of trials, and across many random seeds, all leading to the same conclusion. The only thing that sometimes changed was the number of trials surrounding the learning point that show significantly elevated VTE rates.

5.5 References

Becker, J. T., Olton, D. S., Anderson, C. A., & Breitinger, E. R. P. (1981). Cognitive mapping in rats: The role of the hippocampal and frontal systems in retention and reversal. *Behavioural Brain Research*, *3*(1), 1–22. https://doi.org/10.1016/0166-4328(81)90025-5

Benchenane, K., Peyrache, A., Khamassi, M., Tierney, P. L., Gioanni, Y., Battaglia, F. P., & Wiener, S. I. (2010). Coherent Theta Oscillations and Reorganization of Spike Timing in the Hippocampal-Prefrontal Network upon Learning. *Neuron*, *66*(6), 921–936. https://doi.org/10.1016/j.neuron.2010.05.013

Bett, D., Allison, E., Murdoch, L. H., Kaefer, K., Wood, E. R., & Dudchenko, P. A. (2012). The neural substrates of deliberative decision making: Contrasting effects of hippocampus lesions on performance and vicarious trial-and-error behavior in a spatial memory task and a visual discrimination task. *Frontiers in Behavioral Neuroscience*, *6*. https://doi.org/10.3389/fnbeh.2012.00070

Bett, D., Murdoch, L. H., Wood, E. R., & Dudchenko, P. A. (2015). Hippocampus, delay discounting, and vicarious trial-and-error. *Hippocampus*, *25*(5), 643–654. https://doi.org/10.1002/hipo.22400

Blumenthal, A., Steiner, A., Seeland, K., & David Redish, A. (2011). Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task. *Neurobiology of Learning and Memory*, *95*(3), 376–384. https://doi.org/10.1016/j.nlm.2011.01.011

Brown, V. J., & Tait, D. S. (2010). Behavioral Flexibility: Attentional Shifting, Rule Switching and Response Reversal. In I. P. Stolerman (Ed.), *Encyclopedia of Psychopharmacology* (pp. 209–213). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68706-1_340

Butter, C. M. (1969). Perseveration in extinction and in discrimination reversal tasks following selective frontal ablations in Macaca mulatta. *Physiology & Behavior*, *4*(2), 163–171. https://doi.org/10.1016/0031-9384(69)90075-4

Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal executive and cognitive functions in rodents: Neural and neurochemical substrates. *Neuroscience & Biobehavioral Reviews*, *28*(7), 771–784. https://doi.org/10.1016/j.neubiorev.2004.09.006

Floresco, S. B., Seamans, J. K., & Phillips, A. G. (1997). Selective Roles for Hippocampal, Prefrontal Cortical, and Ventral Striatal Circuits in Radial-Arm Maze Tasks With or Without a Delay. *Journal of Neuroscience*, *17*(5), 1880–1890. https://doi.org/10.1523/JNEUROSCI.17-05-01880.1997

Gentry, E. (1930). *A substitution for trial and error in the white rat* [Master's Thesis]. University of Colorado, Boulder.

George, A. E., Stout, J. J., & Griffin, A. L. (2023). Pausing and reorienting behaviors enhance the performance of a spatial working memory task. *Behavioural Brain Research*, *446*, 114410. https://doi.org/10.1016/j.bbr.2023.114410

Goldman-Rakic, P. S. (1991). Chapter 16 Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. In H. B. M. Uylings, C. G. Van Eden, J. P. C. De Bruin, M. A. Corner, & M. G. P. Feenstra (Eds.), *Progress in Brain Research* (Vol. 85, pp. 325–336). Elsevier. https://doi.org/10.1016/S0079-6123(08)62688-6

Goss, A., & Wischner, G. (1956). Vicarious Trial and Error and Related Behavior. *Psychological Bulletin*, *53*(1), 20.

Goudar, V., Kim, J.-W., Liu, Y., Dede, A. J. O., Jutras, M. J., Skelin, I., Ruvalcaba, M., Chang, W., Fairhall, A. L., Lin, J. J., Knight, R. T., Buffalo, E. A., & Wang, X.-J. (2023). *Comparing rapid rule-learning strategies in humans and monkeys* (p. 2023.01.10.523416). bioRxiv. https://doi.org/10.1101/2023.01.10.523416

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404–411. https://doi.org/10.1037/h0059831

Hallock, H. L., Wang, A., & Griffin, A. L. (2016). Ventral Midline Thalamus Is Critical for Hippocampal–Prefrontal Synchrony and Spatial Working Memory. *Journal of Neuroscience*, *36*(32), 8372–8389. https://doi.org/10.1523/JNEUROSCI.0991-16.2016

Hasz, B. M., & Redish, A. D. (2020). Spatial encoding in dorsomedial prefrontal cortex and hippocampus is related during deliberation. *Hippocampus*, *30*(11), 1194–1208. https://doi.org/10.1002/hipo.23250

Hones, V. I., & Mizumori, S. J. Y. (2022). Response Flexibility: The Role of the Lateral Habenula. *Frontiers in Behavioral Neuroscience*, *16*. https://www.frontiersin.org/articles/10.3389/fnbeh.2022.852235

Horst, N. K., & Laubach, M. (2012). Working with memory: Evidence for a role for the medial prefrontal cortex in performance monitoring during spatial delayed alternation. *Journal of Neurophysiology*, *108*(12), 3276–3288. https://doi.org/10.1152/jn.01192.2011

Hu, D., & Amsel, A. (1995). A simple test of the vicarious trial-and-error hypothesis of hippocampal function. *Proceedings of the National Academy of Sciences*, *92*(12), 5506–5509. https://doi.org/10.1073/pnas.92.12.5506

Hyman, J., Zilli, E., Paley, A., & Hasselmo, M. (2010). Working memory performance correlates with prefrontal-hippocampal theta interactions but not with prefrontal neuron firing rates. *Frontiers in Integrative Neuroscience*, *4*. https://www.frontiersin.org/articles/10.3389/neuro.07.002.2010

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, *345*, 12–26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Izquierdo, A., & Jentsch, J. D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology*, *219*(2), 607–620. https://doi.org/10.1007/s00213-011-2579-7

Jackson, L. L. (1943). VTE on an elevated T-maze. *Journal of Comparative Psychology*, *36*(2), 99–107. https://doi.org/10.1037/h0058536

Johnson, A., & Redish, A. D. (2007). Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point. *Journal of Neuroscience*, *27*(45), 12176–12189. https://doi.org/10.1523/JNEUROSCI.3761-07.2007

Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2015). Vacillation, indecision and hesitation in moment-by-moment decoding of monkey motor cortex. *ELife*, *4*, e04677. https://doi.org/10.7554/eLife.04677

Kay, K., Chung, J. E., Sosa, M., Schor, J. S., Karlsson, M. P., Larkin, M. C., Liu, D. F., & Frank, L. M. (2020). Constant Sub-second Cycling between Representations of Possible Futures in the Hippocampus. *Cell*, *180*(3), 552-567.e25. https://doi.org/10.1016/j.cell.2020.01.014

Kemble, E. D., & Beckman, G. J. (1970). Vicarious trial and error following amygdaloid lesions in rats. *Neuropsychologia*, *8*(2), 161–169. https://doi.org/10.1016/0028-3932(70)90003-5

Kidder, K. S., Gillis, R., Miles, J. T., & Mizumori, S. (2023). *The medial prefrontal cortex during flexible decisions: Evidence for its role in distinct working memory processes* (p. 2023.05.22.541807). bioRxiv. https://doi.org/10.1101/2023.05.22.541807

Kidder, K. S., Miles, J. T., Baker, P. M., Hones, V. I., Gire, D. H., & Mizumori, S. J. Y. (2021). A selective role for the mPFC during choice and deliberation, but not spatial memory retention over short delays. *Hippocampus*, hipo.23306. https://doi.org/10.1002/hipo.23306

Kolb, B., Nonneman, A. J., & Singh, R. K. (1974). Double dissociation of spatial impairments and perseveration following selective prefrontal lesions in rats. *Journal of Comparative and Physiological Psychology*, *87*(4), 772–780. https://doi.org/10.1037/h0036970

Kreher, M. A., Johnson, S. A., Mizell, J.-M., Chetram, D. K., Guenther, D. T., Lovett, S. D., Setlow, B., Bizon, J. L., Burke, S. N., & Maurer, A. P. (2019). The perirhinal cortex supports spatial intertemporal choice stability. *Neurobiology of Learning and Memory*, *162*, 36–46. https://doi.org/10.1016/j.nlm.2019.05.002

Maggi, S., Hock, R. M., O'Neill, M., Buckley, M. J., Moran, P. M., Bast, T., Sami, M., & Humphries, M. D. (2022). *Tracking subject's strategies in behavioural choice experiments at trial resolution* (p. 2022.08.30.505807). bioRxiv. https://doi.org/10.1101/2022.08.30.505807

Mahut, H. (1971). Spatial and object reversal learning in monkeys with partial temporal lobe ablations. *Neuropsychologia*, *9*(4), 409–424. https://doi.org/10.1016/0028-3932(71)90005-4

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), Article 9. https://doi.org/10.1038/s41593-018-0209-y

McLaughlin, A. E., & Redish, A. D. (2023). Optogenetic disruption of the prelimbic cortex alters long-term decision strategy but not valuation on a spatial delay discounting task. *Neurobiology of Learning and Memory*, *200*, 107734. https://doi.org/10.1016/j.nlm.2023.107734

Medin, D. L., Borkhius, M. L., & David, R. T. (1970). Response latency and brightness judgments by monkeys. *Journal of Experimental Psychology*, *83*(3, Pt.1), 480–485. https://doi.org/10.1037/h0028875

Meyer-Mueller, C., Jacob, P.-Y., Montenay, J.-Y., Poitreau, J., Poucet, B., & Chaillan, F. A. (2020). Dorsal, but not ventral, hippocampal inactivation alters deliberation in rats. *Behavioural Brain Research*, *390*, 112622. https://doi.org/10.1016/j.bbr.2020.112622

Miles, J. T., Kidder, K. S., Wang, Z., Zhu, Y., Gire, D. H., & Mizumori, S. J. Y. (2021). A Machine Learning Approach for Detecting Vicarious Trial and Error Behaviors. *Frontiers in Neuroscience*, *15*. https://www.frontiersin.org/article/10.3389/fnins.2021.676779

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moore, T. L., Killiany, R. J., Herndon, J. G., Rosene, D. L., & Moss, M. B. (2005). A non-human primate test of abstraction and set shifting: An automated adaptation of the Wisconsin Card Sorting Test. *Journal of Neuroscience Methods*, *146*(2), 165–173. https://doi.org/10.1016/j.jneumeth.2005.02.005

Muenzinger, K. F. (1938). Vicarious Trial and Error at a Point of Choice: I. A General Survey of its Relation to Learning Efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology*, *53*(1), 75–86. https://doi.org/10.1080/08856559.1938.10533799

Muenzinger, K. F., & Gentry, E. (1931). Tone discrimination in white rats. *Journal of Comparative Psychology*, *12*(2), 195–206. https://doi.org/10.1037/h0072238

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols*, *14*(7), Article 7. https://doi.org/10.1038/s41596-019-0176-0

Papale, A. E., Stott, J. J., Powell, N. J., Regier, P. S., & Redish, A. D. (2012). Interactions between deliberation and delay-discounting in rats. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(3), 513–526. https://doi.org/10.3758/s13415-012-0097-7

Papale, A. E., Zielinski, M. C., Frank, L. M., Jadhav, S. P., & Redish, A. D. (2016). Interplay between Hippocampal Sharp-Wave-Ripple Events and Vicarious Trial and Error Behaviors in Decision Making. *Neuron*, *92*(5), 975–982. https://doi.org/10.1016/j.neuron.2016.10.028

Ragozzino, M. E. (2007). The Contribution of the Medial Prefrontal Cortex, Orbitofrontal Cortex, and Dorsomedial Striatum to Behavioral Flexibility. *Annals of the New York Academy of Sciences*, *1121*(1), 355–375. https://doi.org/10.1196/annals.1401.013

Ragozzino, M. E., Kim, J., Hassert, D., Minniti, N., & Kiang, C. (2003). The contribution of the rat prelimbic-infralimbic areas to different forms of task switching. *Behavioral Neuroscience*, *117*(5), 1054–1065. https://doi.org/10.1037/0735-7044.117.5.1054

Redish, A. D. (2016). Vicarious trial and error. *Nature Reviews Neuroscience*, *17*(3), 147–159. https://doi.org/10.1038/nrn.2015.30

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, *461*(7261), Article 7261. https://doi.org/10.1038/nature08275

Roberts, A. C., Robbins, T. W., & Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *The Quarterly Journal of Experimental Psychology Section B*, *40*(4), 321–341. https://doi.org/10.1080/14640748808402328

Santos-Pata, D., & Verschure, P. F. M. J. (2018). Human Vicarious Trial and Error Is Predictive of Spatial Navigation Performance. *Frontiers in Behavioral Neuroscience*, *12*. https://doi.org/10.3389/fnbeh.2018.00237

Saravanan, V., Berman, G. J., & Sober, S. J. (2020). Application of the hierarchical bootstrap to multi-level data in neuroscience. *Neurons, Behavior, Data Analysis and Theory*, *3*(5), https://nbdt.scholasticahq.com/article/13927-application-of-the-hierarchical-bootstrap-to-multi-level-data-in-neuroscience.

Schmidt, B., Duin, A. A., & Redish, A. D. (2019). Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. *Journal of Neurophysiology*, *121*(6), 1981–2000. https://doi.org/10.1152/jn.00793.2018

Schmidt, B., Papale, A., Redish, A. D., & Markus, E. J. (2013). Conflict between place and response navigation strategies: Effects on vicarious trial and error (VTE) behaviors. *Learning & Memory*, *20*(3), 130–138. https://doi.org/10.1101/lm.028753.112

Steiner, A. P., & Redish, A. D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nature Neuroscience*, *17*(7), Article 7. https://doi.org/10.1038/nn.3740

Stout, J. J., George, A. E., Kim, S., Hallock, H. L., & Griffin, A. L. (2023). *Harnessing prefrontal-hippocampal theta synchrony to enhance memory-guided choice* (p. 2023.04.02.535279). bioRxiv. https://doi.org/10.1101/2023.04.02.535279

Stout, J. J., Hallock, H. L., George, A. E., Adiraju, S. S., & Griffin, A. L. (2022). The ventral midline thalamus coordinates prefrontal–hippocampal neural synchrony during vicarious trial and error. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-14707-8

Tang, W., Shin, J. D., & Jadhav, S. P. (2021). Multiple time-scales of decision-making in the hippocampus and prefrontal cortex. *ELife*, *10*, e66227. https://doi.org/10.7554/eLife.66227

Tolman, E. C. (1926). A behavioristic theory of ideas. *Psychological Review*, *33*(5), 352–369. https://doi.org/10.1037/h007053

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. https://doi.org/10.1037/h0061626

Uddin, L. Q. (2021). Cognitive and behavioural flexibility: Neural mechanisms and clinical considerations. *Nature Reviews Neuroscience*, *22*(3), Article 3. https://doi.org/10.1038/s41583-021-00428-w

Voss, J. L., & Cohen, N. J. (2017). Hippocampal-cortical contributions to strategic exploration during perceptual discrimination. *Hippocampus*, *27*(6), 642–652. https://doi.org/10.1002/hipo.22719

Wang, G.-W., & Cai, J.-X. (2006). Disconnection of the hippocampal–prefrontal cortical circuits impairs spatial working memory performance in rats. *Behavioural Brain Research*, *175*(2), 329–336. https://doi.org/10.1016/j.bbr.2006.09.002

Yang, Y., & Mailman, R. B. (2018). Strategic neuronal encoding in medial prefrontal cortex of spatial working memory in the T-maze. *Behavioural Brain Research*, *343*, 50–60. https://doi.org/10.1016/j.bbr.2018.01.020

Yoon, T., Okada, J., Jung, M. W., & Kim, J. J. (2008). Prefrontal cortex and hippocampus subserve different components of working memory in rats. *Learning & Memory*, *15*(3), 97–105. https://doi.org/10.1101/lm.850808

Zielinski, M. C., Shin, J. D., & Jadhav, S. P. (2019). Coherent Coding of Spatial Position Mediated by Theta Oscillations in the Hippocampus and Prefrontal Cortex. *Journal of Neuroscience*, *39*(23), 4550–4565. https://doi.org/10.1523/JNEUROSCI.0106-19.2019

Chapter 6

Conclusion

The main conclusion of this thesis is that the contexts in which behaviors happen are important for helping us interpret those behaviors and their neural underpinnings. We use a variation in decision-making behavior known as vicarious trial and error (VTE) to demonstrate that VTE may reflect different underlying cognitive processes depending on the context in which it is carried out. In particular, we propose that VTE indicates deliberation during periods when tasks are done proficiently and flexibly, while VTE preceding errors and during inflexible periods is more likely a sign of indecision or uncertainty. It's possible that this difference may also involve different neural circuitry or systems, and, as such, could partially explain our imperfect ability to identify VTE from hippocampal oscillations.

In addition, we provide further evidence that the hippocampus is likely only one member of a larger neural system involved in supporting VTE behavior. Part of this system almost certainly includes the medial prefrontal cortex (mPFC). By optogenetically disrupting the mPFC, we can decrease VTE rates during simple decision-making tasks with relatively low working memory requirements, or decouple VTE from changing task dynamics in tasks with additional working memory and behavioral flexibility demands. In the former circumstance, we also see correlations between decreased VTE rates and decreased accuracy, suggesting that disruption impairs the typical deliberative decision-updating that often leads to correct choices. In the latter circumstance, disruption does not decrease VTE prevalence, but does prevent the typical pattern of aggregation at the beginning of blocks followed by suppression toward the end of blocks. This could indicate that the mPFC is not only important for successful updating and retrieval of items in working memory, but also that it typically enables uncertainty assessments, all of which combine to support flexible behavior adjustments.

Importantly, VTE is likely one of the many behaviors with similar motor patterns or trajectories through space that could be utilized by an animal in different ways depending on behavioral context. Raising a hand in a classroom can indicate a desire to be called upon if it's done after a question whereas raising a hand while attendance is taken is done to indicate your presence. Similarly, waving a hand when you first see someone typically indicates a greeting, but signals a departure when done before leaving. In both circumstances, disambiguating the utility or function of the behavior requires an understanding of how an agent interacts with its environment and could not be done by focusing on the agent alone. The point being, kinematics and identification of behavioral motifs alone will not always fully characterize behaviors. By extension, we should not expect a similar *looking* behavior to have similar neural correlates in different contexts.

Another point worth making explicit is that, while much of this thesis has focused on a reduction of cognitive processes into component or dissociable parts, it's clear that many of these processes depend on one another. While working memory may not rely on decision-making, decision-making often relies on working memory. Even "distinct" working memory processes, such as encoding, maintenance, and retrieval often rely on one another, although each subprocess may have its own fault points or mechanisms. For example, maintenance can only be carried out when either encoding or retrieval has occurred, but a startle or attentional redirection during the maintenance period could alter or abolish the representation being maintained. Likewise, poor fidelity encoding of recent information could appear as poor retrieval if that information is required for upcoming behavior.

Hopefully this thesis helps highlight that forming or having cognitive maps is decidedly different from using them. Even if the primary function of the maps themselves is the transformation of related items into structured, neural encodings, access to this information endows an agent with the ability to perform operations on, or with respect to, these maps. In other words, constructing maps is an important prerequisite for their use, of course, but how they are used is what guides or enables certain behaviors and cognitive processes. Is learning analogous to creating the maps, or charting new courses through them? Perhaps it depends on whether some form of map can be recalled from memory. Is decision-making the process of comparing and selecting routes, planning paths, or selecting a particular map? We may find that none of these things are mutually exclusive, but that the answer depends heavily on the context.