

Information Processing in the Orbitofrontal Cortex and the  
Ventral Striatum in Rats Performing an Economic  
Decision-Making Task

A DISSERTATION  
SUBMITTED TO THE FACULTY OF  
THE UNIVERSITY OF MINNESOTA  
BY

Jeffrey Joseph Stott

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

A. David Redish, Adviser

September, 2015

© Jeffrey Joseph Stott, 2015  
ALL RIGHTS RESERVED

## Acknowledgements

I want to thank everyone in the Redish Lab. Without them, this research would not have been possible. I would like to express my gratitude to Dr. David Redish. He has been an excellent mentor and is a creative scientist. His hard work and dedication to the lab is deeply appreciated. I thank him for taking me on as his student. I would like to thank my committee for their help and feedback: Dr. Matthew Chafee, Dr. Mark Thomas, and Dr. Jonathan Gewirtz.

I would like to extend a special thanks to Dr. Virginia Seybold and to the late Sabina Deressa, who both encouraged me at a time when I especially needed it. Thanks also to Dr. Colin Campbell and Evelyn Juliussen. Thank you to John Paton for patiently helping with everything administrative, and to Lateph Onikoro, for computer support.

Most of all, I would like to thank my Mom, my Dad, my sisters Jenny and Janine, and my brother Mike. I could not have done it without them. Thanks also to the Cuffeys for their friendship. Thank you to Diana for her love and constant support. Thank you to the Wallins, who have treated me like family. I would also like to thank the GPN faculty, with a special nod to Chris Honda for his sincerity and personal attention. I received an excellent neuroscience education at the University of Minnesota, and I hope to put it to good use in my future work.

The work in this thesis was supported by NIH grants T32-GM-008471 and R01-DA-030672.

## Abstract

The orbitofrontal cortex (OFC) and ventral striatum (vStr) are key brain structures that represent information about value during decision-making tasks. Despite their very different anatomical properties, numerous studies have found similar patterns of value-related signaling in these structures. In particular, both structures are intimately involved in delay-discounting tasks, which involve a tradeoff between reward magnitude and delay to reward. However, the overlapping activity profiles of these brain regions makes it difficult to tease apart their specific contributions to delay-discounting behavior, and to economic decision-making more generally. In order to better understand the contributions of these two regions to value-based choice, we made simultaneous recordings in the OFC and vStr in rats performing a spatial variant of a traditional delay-discounting task. This allowed us to compare OFC and vStr activity directly in the same subjects while they engaged in a prototypical economic decision-making task, and additionally it allowed us to leverage the tools of spatial decoding analysis to measure non-local reward signaling.

Chapter 1 provides an introduction to current theories of OFC and vStr function within the decision-making literature, in particular contrasting the concepts of neuroeconomics with the multiple decision-making systems framework. Chapter 2 describes the methods used in this thesis, including the design of the spatial delay-discounting task and the analysis of the neural data. Chapter 3 presents the results of single-unit and Bayesian decoding analyses from this dataset. We found that activity in the OFC and vStr was quite similar at the single-unit level, and inconsistent with the neuroeconomic account of value signaling in a common currency. Instead, when we looked specifically at moments of deliberative decision-making (as emphasized by the multiple systems

account), we found important differences between the OFC and vStr. Both the OFC and the vStr showed covert reward signaling during deliberative, vicarious trial-and-error (VTE) behaviors. But vStr signals emerged earlier, before the moment of choice, while covert reward coding in the OFC appeared after the rats had committed to their decision.

These analyses were extended to the level of local field potentials (LFPs), recorded from the same dataset. Local field potentials are a useful tool for studying local processing and interactions between brain regions. Chapter 4 describes the LFP results. Important among these was the finding that the vStr led the OFC at the LFP level (again showing temporal precedence), and furthermore, that the vStr was a stronger driver of OFC activity than vice versa, particularly during VTE. The implications of these results, along with those from the single-unit and Bayesian decoding analyses, are discussed in Chapter 5. Emphasis is placed on our emerging understanding of the role of the vStr in flexible behavior, and how the OFC and the vStr might cooperate to influence value-based choice.

# Table of Contents

Acknowledgements .....	i
Abstract.....	ii
List of Tables .....	v
List of Figures.....	vi
Chapter 1: Introduction.....	1
Chapter 2: Methods.....	15
Chapter 3: Single-Unit and Decoding Analyses .....	44
Chapter 4: Local Field Potential Analyses.....	78
Chapter 5: Discussion .....	104
Bibliography .....	130

# List of Tables

Table 3-1. **Cell counts for delay-responsive cells**..... 60

## List of Figures

Figure 2-1.	<b>The spatial adjusting delay-discounting task .....</b>	<b>17</b>
Figure 2-2.	<b>PSDs from an example session .....</b>	<b>22</b>
Figure 2-3.	<b>Histological determination of recording sites .....</b>	<b>25</b>
Figure 2-4.	<b>Schematic diagram of procedure for calculating z-scores.....</b>	<b>27</b>
Figure 2-5.	<b>Heat/line plot showing raw firing rate data around reward receipt .....</b>	<b>29</b>
Figure 2-6.	<b>Example non-VTE and VTE events .....</b>	<b>32</b>
Figure 2-7.	<b>Example of LFP data “pre-whitening” for Granger causality analysis .....</b>	<b>39</b>
Figure 2-8.	<b>Example of LFP data showing the cross-correlation technique.....</b>	<b>42</b>
Figure 3-1.	<b>Behavior on the delay discounting task reflects distinct phases .....</b>	<b>46</b>
Figure 3-2.	<b>Titration behavior on the spatial adjusting delay-discounting task .....</b>	<b>47</b>
Figure 3-3.	<b>Distribution of VTE events as quantified by path curvature.....</b>	<b>49</b>
Figure 3-4.	<b>Relationship between task phase, lap type, and VTE .....</b>	<b>50</b>
Figure 3-5.	<b>Relationship between VTE events and lap type .....</b>	<b>51</b>
Figure 3-6.	<b>Reward responsive units in vStr and OFC .....</b>	<b>53</b>
Figure 3-7.	<b>Single cells code for reward and for value .....</b>	<b>55</b>
Figure 3-8.	<b>Example “delay-discounting” neuron .....</b>	<b>58</b>
Figure 3-9.	<b>Distribution of delay-coding cells on the DD task .....</b>	<b>59</b>
Figure 3-10.	<b>Correlation between delay coding and reward preference .....</b>	<b>63</b>
Figure 3-11.	<b>Ensemble decoding to the reward sites .....</b>	<b>66</b>
Figure 3-12.	<b>Population firing rate PETHs .....</b>	<b>67</b>
Figure 3-13.	<b>Reward site decoding during VTE .....</b>	<b>68</b>
Figure 3-14.	<b>Separate feeder site representations during choice.....</b>	<b>70</b>
Figure 3-15.	<b>Spatial control for Figure 3-13.....</b>	<b>72</b>
Figure 3-16.	<b>Spatial control for Figure 3-14.....</b>	<b>73</b>

Figure 4-1.	<b>Spectral peaks in recorded areas .....</b>	<b>81</b>
Figure 4-2.	<b>Example oscillatory activity in the recorded brain areas .....</b>	<b>82</b>
Figure 4-3.	<b>Self-coherence plots for OFC, vStr, and HIPPP .....</b>	<b>83</b>
Figure 4-4.	<b>Example power-frequency spectrograms and coherogram between OFC and vStr .....</b>	<b>84</b>
Figure 4-5.	<b>Gamma power at the choice point .....</b>	<b>85</b>
Figure 4-6.	<b>Gamma power around the time of reward receipt.....</b>	<b>87</b>
Figure 4-7.	<b>Average coherence spectra across animals .....</b>	<b>90</b>
Figure 4-8.	<b>OFC-vStr coherence at the choice point .....</b>	<b>91</b>
Figure 4-9.	<b>OFC-vStr coherence at the time of reward receipt.....</b>	<b>92</b>
Figure 4-10.	<b>Average Granger causality between OFC and vStr.....</b>	<b>94</b>
Figure 4-11.	<b>Granger causality at the choice point and during reward receipt.....</b>	<b>96</b>
Figure 4-12.	<b>Maximum cross-correlation distributions .....</b>	<b>98</b>
Figure 5-1.	<b>Schematic illustration of parallel neuroeconomic tasks in monkey and rat.....</b>	<b>119</b>

# Chapter 1: Introduction

## The neuroeconomic model of choice

Decision-making, in the simplest sense, involves choosing between different options. We face countless decisions in our everyday lives, and these decisions can vary along any number of dimensions. For example, when choosing where to eat out for lunch, one is likely to consider the prices of the different options (cost), the waiting times at different restaurants (delay), one's fondness for different types of food (preference), and possibly more abstract considerations, such as long term health goals (nutritional content). The mechanism by which different, often opposing attributes of an item factor into an agent's<sup>1</sup> choice is a central question for the field of decision-making research.

Across different disciplines [1-3], the idea that binds these disparate attributes together is the concept of "value." Value, broadly defined, expresses how desirable an object or outcome is to an agent. Economic theories in particular assert as an axiom that agents should integrate all of the relevant decision variables into a single unitary value construct [4, 5]. This theoretical framework has the advantage of solving (or perhaps, sidestepping) two difficult problems: (1) how value is calculated when different factors affect the outcome, and (2) how different outcomes are compared.

To the first point, this axiom provides a means by which the influence of different economic variables (reward magnitude, probability, delay, etc.) can be combined. Each economic variable is related to value by a utility function. For instance, a \$100 gift card is worth \$100 if it is immediately available, but if it can only be redeemed six months later, its value (considered in the present moment) is reduced. In practice, determining the utility function can be difficult. Value is inherently subjective, and it can only be quantified by making inferences based on the behavior of the

---

<sup>1</sup> The term agent, as used here, encompasses humans, animals, and decision-making algorithms.

individual. Conceptually however, these different “decision variables” can be weighted together—integrating the costs and benefits—into a single number.

To the second point, regarding the comparison of outcomes, value representations are deemed to be expressed in a “common currency.” In the same way that money serves as a common currency in the marketplace—allowing for the exchange of goods—valuation by the individual agent in a common *internal* currency allows for the direct comparison of even qualitatively very different options. Which good or option to choose becomes a value maximization problem. This framework has intuitive appeal. It matches our intuitive notion that when we make decisions, we weigh them against each other (integrating costs and benefits). Mathematically, the common currency principle, among other assumptions, is necessary for internally consistent (i.e. rational) choices. Many economists are quick to point out that their models describe decision-making “as if” the agent were applying the steps in the model. They make no claims to the actual cognitive or neural events taking place in the human (or animal) brain [6]. Others have taken the stance that mathematical models need to explain and make predictions about cognitive operations if they are to have true validity and practical relevance [7]. From either perspective, the common currency framework greatly simplifies the process of modeling decision-making behavior.

Decisions arise from computations occurring in neural systems. Under the assumption that value is calculated in a common currency, one would expect to find abstract value representations somewhere in the nervous system. Therefore, it generated a great deal of excitement when neuroscientists studying decision-making started identifying brain areas with neural correlates of value (see [8, 9] for a historical perspective on these developments), including the orbitofrontal cortex (OFC) and the ventral striatum (vStr).

### Value signals in the OFC and vStr

My thesis is motivated by the question of how the OFC and the vStr contribute to value-based decision-making, particularly during deliberative behavior. Two lines of evidence in

particular suggest that the OFC and the vStr—particularly the nucleus accumbens (NAc) region—could be candidate areas for calculating value. The first comes from functional magnetic resonance imaging (fMRI) studies in human subjects. fMRI studies have repeatedly found that the OFC and vStr are engaged during the anticipation of reward [10-13], and that both structures show activity that scales with the expected value (EV) of reward when subjects are offered the choice between differently valued reward options [12, 14, 15]. See [8] and [16] for meta-analyses of fMRI studies linking subjective value representations to the OFC and vStr.

The second line of evidence comes from recording studies in the monkey OFC. For example, Padoa-Schioppa and colleagues gave monkeys differently flavored juice options, and measured the monkeys' preferences along a common scale (as a hypothetical example, 3 units of grape Kool-Aid = 1 unit of apple juice). They discovered neurons that changed their firing rate as a function of the behaviorally-inferred subjective value of the juice offers [17-19]. Importantly, this neural activity was independent of (or “abstracted from”) the sensory and motor variables of the task [20]. Some recording experiments from different research groups support this interpretation, in that the OFC shows strong coding of decision variables (reward magnitude, delay to reward, reward preference, satiety), largely without spatial tuning [21-27]. However, in studies that have manipulated more than one decision variable on the same task (e.g. delay + effort), there has been little evidence for an integrated value signal in single OFC neurons [21, 22, 28](but see [25, 29] for counterexamples).

In general, fMRI studies favor the interpretation of abstract value signals in the ventromedial prefrontal cortex (vmPFC)/OFC [8, 30-32]<sup>2</sup>, although some authors have highlighted

---

<sup>2</sup> Interestingly, authors on the whole seem to emphasize the role for the OFC, even when both the OFC and vStr are implicated. In the review by Levy & Glimcher [8], both NAc and OFC were consistently identified with value coding, according to their meta-analysis. But the nucleus accumbens does not appear in the abstract or feature in their discussion. The authors argue that vStr activity is complicated by the fact that (1) subjective value signals have been found in the dorsal striatum in single unit studies in the monkey (by that logic, OFC activity is complicated by the fact that subjective value signals have been seen in many different prefrontal cortical areas in single unit studies [33]) and (2) vStr signals are often associated with learning, rather than value *per se*. Again, the same thing could be said for the OFC. Most likely, the authors

the relative nature of OFC value signaling and the coding of sensory information in fMRI experiments [34]. Regardless, the study by Padoa-Schioppa & Assad [17] and the evidence from fMRI studies, taken as a whole, have driven the idea that the OFC in particular calculates value in an abstract, common neural currency [35]—the neuroeconomic model of OFC function.

In contrast, neuroeconomic experiments in rodents have largely failed to find these representations of abstract value signaling in the OFC [36, 37]. Instead, rodent recording studies have found that the OFC encodes value-related associations, but in the context of different internal and external task variables. Examples include spatial tuning [37, 38], stimulus identity [39, 40], and decision confidence [41], each of which are value-neutral and irrelevant to a purely economic value signal. Various, other studies have reported that the OFC can also encode information about the animal's past history of choices [42], chosen value signals [42], task schemas [43], and even regret [44]. Together, these data support the notion that the rodent OFC generates state-based representations of reward, combining pure reward information with information about the structure of the task [45]. These findings contradict the neuroeconomic model of OFC function.

Moreover, rodent studies have identified a prominent role for value coding in the ventral striatum (vStr) [44, 46-50], calling into question a *unique* role for the OFC in signaling value. Taken together, the data paint a mixed portrait of the roles of the OFC and vStr in signaling value. These two brain areas show a high degree of similarity in their activity as measured with fMRI [8, 13, 16]. Neural recording studies also show overlapping activity between OFC and vStr [51]. In rodents and monkeys, the OFC and vStr both respond to reward-related cues and to reward receipt [52-54], and neural firing in these structures can be modulated by sensory and motor information [24, 38, 49, 55]. Thus, it remains to be determined exactly how their roles differ during value-based decision-making.

---

are referring to reward prediction errors (RPEs) in the vStr seen in the BOLD signal, which is almost certainly due to afferent dopaminergic inputs from the VTA/SNc. RPEs are not seen in the vStr with electrophysiological techniques.

## Multiple decision-making systems

The preceding discussion has treated value-based decision making as a single process. In reality, our decision-making strategies can vary widely depending on the type of decision and the context. A classic example is driving to work. If you have moved to a new city and you are first learning your way to work, you are likely to pay attention to street names, distal landmarks, and the (compass) direction you are heading. You are building a mental map of the neighborhood to estimate your position on the map and navigate through it. Once you have driven to work many, you can drive “on autopilot,” going straight or turning at each stoplight in response to familiar stimuli. You no longer need the mental map to accomplish the task of driving to work. Thus, the same overt behavior can be realized by different cognitive operations. The same holds true for value-based decision-making. Performing or “solving” a task can be accomplished by different action-selection systems, with different underlying value representations, possibly mediated by different brain areas.

The converse also holds true. The same brain structure can be implicated in numerous behaviors, often very different from one another. The OFC and the vStr have both been linked to simple forms of Pavlovian conditioning which do not require an instrumental response and which do not necessitate a representation of the outcome (i.e. the unconditioned stimulus) [16, 56]. These two structures have also been linked to goal-directed instrumental behavior, which requires an explicit representation of the potential outcome [57, 58]. Measurements of firing rates or BOLD activity on both kinds of task reveal neural correlates of value. Therefore, it is essential to distinguish which decision-making processes are being engaged when interpreting these signals.

Convergent evidence supports the idea that there are multiple<sup>3</sup> decision-making systems in the mammalian brain, which employ distinct computational mechanisms [63, 65, 66], and which

---

<sup>3</sup> The exact number of systems and the precise terminology depends on the author. However, there is a wide degree of consensus about the general framework. In particular, the distinction between habitual and

have distinct (but sometimes overlapping) anatomical loci: (1) A hardwired system of reflexes that are automatically evoked by certain stimuli. (2) A Pavlovian system that releases evolutionarily conserved emotional and visceral reactions in response to learning between an initially neutral conditioned stimulus and an appetitive or aversive unconditioned stimulus. (3) A Deliberative system (also called “goal-directed” or “model-based”) that is capable of planning and future-oriented behavior. Goal-directed behavior involves an understanding of the causal relationship between actions and their outcomes, and the control of actions according to an expectation of those outcomes—termed action-outcome (A-O) associations [67]<sup>4</sup>. (4) A Habitual system (also called “model-free”) that executes well-learned behaviors based on stored stimulus-response (S-R) associations.

The habitual and goal-directed systems are thought of more or less as opposites, or perhaps as complements to one another. The deliberative system comes online early, when animals (or humans) are learning a new task, when task contingencies change, or when faced with a novel situation. The deliberative system is flexible, but computationally intensive, because it involves a search through possible states (and their associated outcomes). The habitual system, in contrast, comes online only with extended experience, after the animal has experienced a consistent relationship between action and reward. Habitual behavior is fast and efficient, and requires less attention, but it is relatively inflexible. The habitual system is most closely associated with the dorsolateral (sensorimotor) striatum [68], while the deliberative system is associated with the hippocampus and prefrontal cortex (including the orbitofrontal cortex), and, we would argue, the ventral striatum [69]. How these systems interact (or compete) with one another is not well understood, and remains an important question for the field [61, 70]. However, the general trend is

---

goal-directed modes of action-selection is ubiquitous [59-62], and the inclusion of a Pavlovian system is generally used [63, 64].

<sup>4</sup> The term “model-based” derives from the reinforcement learning literature [2]. In this nomenclature, goal-directed decisions involve an encoding of the different states of the world (e.g. the task structure), a mapping of the transition probabilities between states and their outcomes, and an explicit representation of the outcomes obtained (including stimulus identity).

robust: flexible decision-making is engaged early, during initial learning, or “exploratory behavior” [64, 71-74]. Over time, behavioral control transitions to a habitual decision-making mode, matching the transition to skilled behavior and the “exploitation” of a previously rewarded strategy [71, 75].

### Multiple kinds of value signaling

Value is represented in these systems in different ways. For the deliberative system, reward valuation should take place *when* and *where* there is need for the intentional consideration of options. For example, during early laps on a T-maze task, reward signals were present in the vStr when rats paused at the final choice point. These reward signals were not present on a similar T-junction on the navigation sequence that did not involve a choice between the two goal arms [48]. Thus, reward signals on deliberative laps showed spatial specificity.

On late laps, once animals had learned the task contingency (turn Left or turn Right) and had automated their behavior (as can be seen on T-maze tasks by increased running speeds, increased path stereotypy, and high, asymptotic levels of performance), reward signaling at the choice point was no longer present. Thus, value signaling that reflects deliberative behavior shows temporal specificity; it occurs when deliberation is likely to prove useful (during early laps, when the rat is uncertain about which rule is in effect—i.e., Left or Right), and it is absent when deliberation is in fact not useful (during late laps, when the rule is understood, and deliberation only wastes time). Value signaling can occur in the absence of any necessary use for action-selection [16]. The result by van der Meer & Redish shows that vStr reward coding at the choice point is consistent with a deliberative process, and not a general predictive process that can be explained by arousal, motivation, salience, or “default” reward prediction. Therefore, deliberative value signals are present at times and places when the animal is uncertain about the best course of action, and a cognitive search process can be used to evaluate different possible outcomes [72, 76].

In contrast to the deliberative system, which calculates value signals based on internally generated expectations of the outcome (A-O associations), the habitual system is thought to use “cached values” of situation-action (S-A) pairs: “If in situation X, perform action Y.” Cached values are learned incrementally—they represent long term estimates of the best course of action—and are slow to change in the face of unexpected consequences. Thus, habit-related value signals should appear relatively late in the learning process and they should be relatively unperturbed by changes in outcome value.

Under the “actor-critic” framework of temporal difference reinforcement learning (TDRL), cached values are instantiated in the synaptic connections between frontal areas that recognize the situation and the dorsolateral striatum (DLS), which store “action propensities” [77]. The dorsolateral striatum is essential for the expression of habits [78], and it has a neurophysiological correlate that could represent cached value [79, 80].

In rats learning a T-maze task, cells in the DLS respond more or less uniformly across the spatial extent of maze during early stages of learning. With extended experience, DLS neurons come to fire selectively at the beginning and end of the maze [79, 81]. This “task-bracketing” activity is thought to reflect ingrained situation-action associations. Interestingly, phasic DLS firing at the beginning of the maze is inversely related to deliberative behaviors seen at the choice point [82]. Moreover, task-bracketing in the DLS is maintained after pairing the rewards with illness, as if it represented an inflexible habit [82]. Thus, “habitual value signals” also appear with spatial and temporal specificity. They evolve slowly over time, matching the timecourse of procedural learning [81], and appear at the start location of movement sequences [82-84], when stimulus-driven action-selection is likely to occur.

#### OFC and vStr in goal-directed versus non-goal-directed behavior

As seen in part in the discussion above, the different action-selection systems are associated with distinct neural structures. The habitual and deliberative action-selection systems have been

most closely associated with the dorsolateral striatum on the one hand, and the hippocampus and prefrontal cortex on the other [66, 69]. The OFC, part of the prefrontal cortex, has been linked with deliberative and goal-directed decision-making [57, 85]. Although the OFC signals stimulus-outcome (S-O) associations generally, some authors [86] have argued—based on lesion data—that the OFC is only *required* for model-based tasks and not for tasks that can be solved with model-free behavior. For example, the OFC is not required for tracking slow fluctuations in value on a probabilistic reward task [87], nor is it required for operant responding to cues with general (vStr. specific) affective value. The OFC is required for outcome-specific Pavlovian to Instrumental transfer (PIT) [88] and reinforcer devaluation [89-91]<sup>5</sup>. This interpretation is satisfying in that it provides a clear-cut categorization. But the ubiquity of reward signaling in the OFC across tasks [53] would appear to make some of this information redundant. It begs the question of what the OFC is doing during those other tasks. Even if the OFC is not strictly necessary for some model-free behaviors, it may contribute to others (it is hard to rule out all cases). In all likelihood, the OFC contributes to both model-based and model-free behavior. However, it is worth noting that contemporary thinking about the OFC emphasizes its role in goal-directed behavior.

The vStr has perhaps a more ambiguous position. In the traditional habitual versus goal-directed dichotomy, the vStr does not fall into either camp [62]. While the dorsolateral striatum is clearly connected to habitual behavior, and the dorsomedial striatum has been associated with the goal-directed system [94], the ventral striatum is thought to modulate behavior across domains by means of general, affective mechanisms. Dopaminergic input to the nucleus accumbens is necessary for the invigoration of behavior [95], including the application of effort to overcome obstacles [96]. The vStr, along with the ventral pallidum, is responsible for attributing “incentive salience” to reward-associated cues, which can drive basic appetitive behavior [97].

---

<sup>5</sup> Note that a role in model-based behavior for the OFC does not imply that it signals value in an abstract, common-currency—the neuroeconomic model of OFC function. These two theoretical accounts are distinct. In fact, the OFC codes sensory specific features of the reward [92], a function that is part of model-based decision-making, but is incompatible with a purely abstract value signal [93].

In line with this general motivational role in behavior, the nucleus accumbens core is necessary for general-affective (i.e. not outcome-specific) Pavlovian to Instrumental transfer [98]. In PIT, the non-contingent presentation of a previously learned appetitive CS is able to increase instrumental responding (i.e. energizing instrumental behavior<sup>6</sup>). Disconnection of the NAc and the basolateral amygdala impairs second-order conditioning [99], a behavior in which the secondary reinforcer influences behavior, irrespective of the current value of the primary reward [100]. The inability to change responding when the value of the outcome has changed is characteristic of model-free behavior (similar to reinforcer devaluation and contingency degradation—standard tests of habitual responding). The NAc is also necessary for conditioned place preference [101] and psychomotor sensitization [102]—important, but simple behaviors that do not qualify as goal-directed. Together, these data are consistent with the proposal of the vStr as a limbic-motor interface [103], influencing motor behavior by way of general affective and motivational processes.

In spite of the evidence that the vStr is chiefly involved in motivational and affective modulation of ongoing behavior, there is emerging evidence that the vStr is also involved in model-based decision-making. The vStr is necessary for reinforcer devaluation [104, 105], a test that measures whether new information about the value of the outcome influences behavior. Outcome-specific PIT is dependent on the nucleus accumbens shell region [98, 106]. As with the OFC, the vStr is also necessary for identity-based unblocking, in which learning is driven by a violation in the expectation of the identity of the reward, with reward value being held constant [92]. This test measures model-based behavior, because an expectation of the specific sensory qualities (i.e. “identity”) of the reward is a signature of model-based behavior [107]. Value-based unblocking, in which the magnitude of reward is changed [108], can be accomplished with a (model-free) scalar value signal.

---

<sup>6</sup> Interestingly, the introduction of the CS has no bearing on the reward contingency in the task. A typical example would be responding on a fixed rate lever-pressing task. Logically, the CS should have no effect on behavior.

Several neurophysiological studies also implicate the vStr in model-based, goal-directed decision-making. The study by van der Meer & Redish [48] is a good example. As discussed above, the vStr exhibited reward representations at the choice point, but only when the rat paused (likely reflecting deliberation), and only during early laps on the task, when deliberation would be expected. In a task that involved either fixed movement sequences or “flexible approach” from different, novel positions, cue-evoked excitation in nucleus accumbens neurons was shown to be necessary for reward-seeking behavior, and dopamine receptor antagonism in the NAc specifically impaired flexible approach trajectories, but not inflexible, stereotyped approach behavior [109, 110]. These lesion and recording data support a role for the vStr in model-based/goal-directed behavior.

#### The spatial adjusting delay-discounting task

As demonstrated by the discussion above, the OFC and the vStr are both implicated in deliberation and goal-directed behavior, and they both show reward-related activity in a variety of situations (model-based and model-free). Given that the same structure can be involved in different types of behavior, and that the same overt behavior can be accomplished by fundamentally different computations (e.g. cached value versus a search and evaluate process), it is important to distinguish which decision-making process is at play when interpreting neural correlates of value. A useful approach in comparing value representations is to record activity on a decision-making task in which the animal or human subject engages in both deliberative and habitual behavior within a given session [111]. Combined with multi-site recordings, this allows for a direct comparison of value signaling between structures, during contrasting decision-making modes.

We designed a task, the spatial adjusting delay-discounting task, that is well-suited for making these comparisons. On this task, rats run laps along a T-maze and choose between an immediate, 1 pellet reward, and a 3 pellet reward with a variable delay (the “adjusting delay”). The spatial delay-discounting task involves a transition from flexible behavior on early laps, as rats

“titrate” the adjusting delay, to habitual behavior on late laps, when rats engage in a routine alternation strategy [112]. Overtraining rats on a simple task is sufficient to induce strong, habitual responding [79, 82]. However, when rats complete this transition within a single day [112], it allows the experimenter to compare neural activity in the same, identified ensemble of cells across conditions (i.e. habitual versus deliberative phases)<sup>7</sup>.

### Vicarious trial and error

Rats on the spatial delay-discounting task show prominent instances of pause-and-look behavior at the choice point [112], called Vicarious Trial and Error (VTE). VTE is thought to be a behavioral manifestation of deliberation in rats [64, 76, 113]. During VTE events, the hippocampus shows non-local decoding that moves ahead of the animal toward the feeder sites in a serial fashion [114]. These transient, forward decoding events are termed “sweeps.” The discovery of sweeps in the hippocampus provided a neural mechanism for future planning in the rodent, as is required by theories of deliberative behavior [115, 116]<sup>8</sup>. It also bolstered the idea that VTE events do indeed reflect deliberation in the rodent. Therefore, VTE provides a useful index of laps in which the rat is likely deliberating. Non-VTE laps, when the rat passes through the choice point rapidly and in a stereotyped trajectory, indicate habitual behavior.

In order to judge the available actions, there needs to be an estimate of the rewards associated with each action. This led to the hypothesis that downstream, reward-responsive regions might be activated by hippocampal activity during sweeps, providing the critical reward evaluation

---

<sup>7</sup> The population of recorded cells can change across days, even when the position of the electrode(s) is held constant. Therefore it is advantageous to have both conditions present within individual recording sessions. This requirement is usually easy to satisfy. However, in the case of habits, early (deliberative) behavior is usually compared to automated behavior many days or weeks later. Having deliberative and habitual phases within the same session therefore has an added technical advantage.

<sup>8</sup> Human imaging studies have shown prospective activity in the cortex and hippocampus during deliberation [117, 118], showing that deliberation in the human also involves the reactivation of mnemonic information (as in sweeps) and is associated with the hippocampus.

component [76]. Subsequent recording studies in the vStr [119] and OFC [120] showed that these two value-related regions displayed increased reward-site decoding during VTE events, providing a possible neural mechanism for reward evaluation during VTE.

#### Relating the OFC and vStr during value-based choice

Comparing the data in these two studies, it appeared that covert reward signals in the vStr might precede those in the OFC. In the case of [48], rats were allowed to “change their mind” and reverse directions after crossing through the choice point. In the study by [120], the timecourse of feeder site decoding was assessed during VTE events at the choice point, similar to the analyses presented here. Therefore, the relative timing of these signals could not be determined with confidence, as these studies recorded from separate sets of animals and used different behavioral measures. The precise timing of activity during the decision-making process can reveal fundamentally distinct neural computations. These considerations made the case for performing simultaneous neural recordings in the OFC and vStr, in order to resolve the temporal dynamics of activity in these regions during deliberative behavior.

Another motivating factor for performing the experiments in this thesis is that a direct comparison of activity in the OFC and vStr (in the same animals, on the same task) would provide useful insights into the respective roles of these two structures during decision-making. The OFC and vStr show similar neural correlates on value-based decision-making tasks. These structures likely work together to influence valuation and choice [52, 86].

Anatomically, the OFC projects to the dorsomedial and ventral striatum, and this cortico-striatal connection is conserved across rodents, monkeys, and humans [66, 121-126]. The striatum, in turn, sends projections back to its cortical inputs via parallel striatal-thalamocortical loops [121, 127]. Gamma oscillations in the 50 Hz range are prominent in both structures [128, 129], which could mediate communication between these brain regions [130] during the choice process. This hypothesis has not been tested.

Functionally, the OFC and the vStr are both intimately involved in delay-discounting behavior, with lesions of either structure reducing (by and large) how long rats are willing to wait for delayed rewards [131-140]. Disconnecting the two regions increases discount rates [141], showing an interaction between structures. Lesioning the OFC eliminates the preferential coding of reward magnitude in vStr neurons [142], showing that the OFC can powerfully shape vStr activity. These examples indicate that the OFC and vStr participate in shared behavioral processes, particularly in the valuation of delayed reward [143], and in flexible behavior [48, 120].

With the above considerations in mind, we performed simultaneous recordings in OFC and vStr on an economic decision-making task—the spatial adjusting delay-discounting task—to better understand the profile and timing of reward-related activity in these two brain structures, especially during deliberative VTE behaviors.

## Chapter 2: Methods

### Data Collection

#### Animals

The experiments described in this thesis were carried out on Fisher Brown Norway rats (Harlan, Indianapolis, IN). Six adult male rats, 7-12 old months at the start of training, were used for the main experiment described in this thesis. Behavioral data from four of these rats were incorporated into the dataset used in Papale et al. [112]. Rats were housed on a 12-hour light/dark cycle and had *ad libitum* access to water in their home cages. Prior to training, rats were handled daily for 2 weeks in order to acclimate them to human contact. In the second week, they were introduced to the experimental food pellets (45mg unflavored food pellets; Research Diets, New Brunswick, NJ, USA). Rats earned their daily food requirement on the maze and were maintained at all times above 80% of their original free-feeding weight. All procedures were conducted in accordance with the National Institutes of Health guidelines for animal care and approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Minnesota.

#### Maze training

Rats were trained on the spatial adjusting delay-discounting T-maze (Figure 2-1), identical to that used in [112]. Rats were first trained to run laps on the task with one choice arm or the other blocked in daily one hour sessions. This “shaping” step was necessary to teach the rats to run toward the feeders for food, and to depart the feeders—starting a new lap—after consuming reward. Once the rats consistently ran 100 laps within the hour, they moved on to the delay-discounting task. For the delay-discounting task, rats were tested in daily one hour sessions, occurring at the same time

each day. Each rat ran a 30-day sequence on the spatial delay-discounting task before surgery in order to thoroughly learn the structure of the task.

The structure of the task was as follows: On each day, one feeder provided a small reward (one food pellet) after one second—the “non-delay side”—while the other feeder provided a large reward (three food pellets) after an adjustable delay—the “delay side.” The left or right position of the delay and non-delay sides changed from session to session, but was counterbalanced across the 30-day sequence. The initial delay was drawn pseudo-randomly from a uniform distribution between 1 and 30 seconds, without replacement. To earn food rewards on the task, rats traversed a navigational sequence from the start of the maze to the choice point and then along one of the choice arms to the reward site, and after receiving reward, back through the start of the maze toward the choice point again (see Figure 2-1).

During performance of the task, the length of the adjusting delay changed based on the behavior of the animal. Successive laps to the delay side increased the adjusting delay by 1 second. Successive laps to the non-delay side decreased the adjusting delay by 1 second. Alternating from side to side kept the adjusting delay constant. Rats were “committed” to their decision once they passed an invisible boundary line that defined the exit from the choice point (and entry into the “feeder zone”—Figure 2-1, orange shaded bars). This invisible line was implemented via the overhead tracking system and with custom written task code in MATLAB (MathWorks, Natick, MA). Once rats crossed over the line, an audible feeder countdown commenced, indicating to the rat that it had entered the feeder zone and made its choice. Additionally, rats were kept from running backwards into the choice point by manually blocking them with a long cardboard tube. Blocking was only necessary during training, and early on during the behavioral sequence. With experience, rats readily learned the position of this boundary and learned not to run backwards, once they had crossed this line. No blocking was necessary for the recording sessions, from which the behavioral and neural data presented in this thesis are taken.

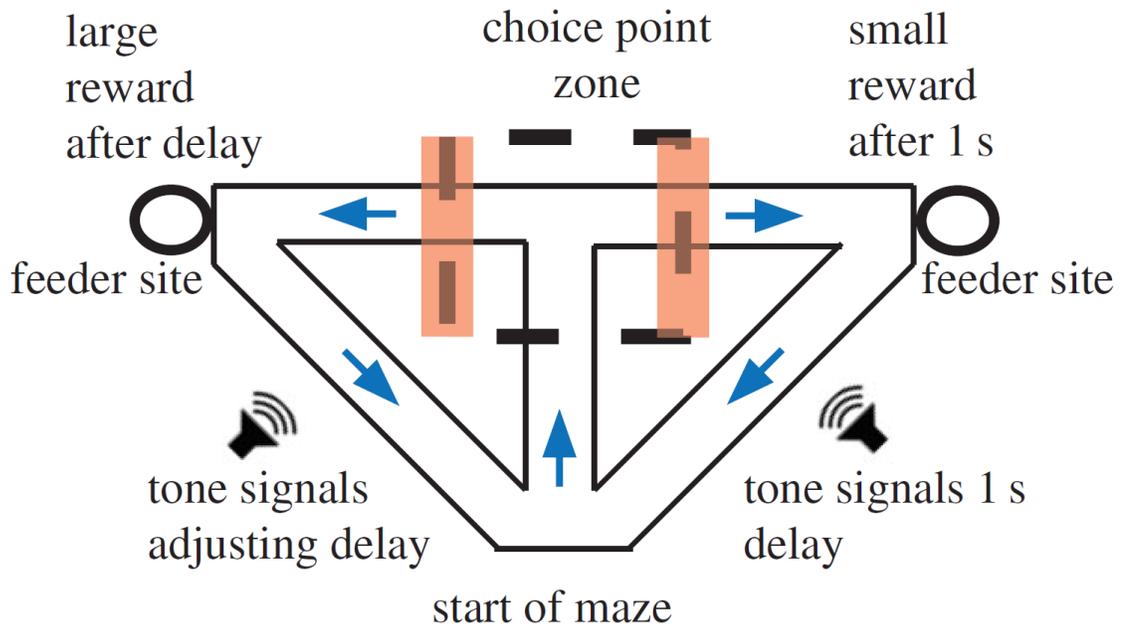


Figure 2-1. **The spatial adjusting delay-discounting task.** At the beginning of each session, rats were placed at the start of the maze. Rats ran through the central stem and into the choice point zone (dashed lines). Tones commenced as soon as rats chose one side by exiting the choice point zone (either on the left or the right, orange bars) and counted down to reward delivery. The small reward feeder site offered 1 pellet after 1 second. The large reward feeder site offered 3 pellets after a variable delay. Figure image from [144], reprinted with permission.

When entering the non-delay side feeder zone, a tone signaled the delivery of reward one second later. When entering the delay side feeder zone, a series of tones descending in pitch provided a countdown to the moment of reward delivery. On each second of the countdown, a 100ms long pure tone was played, descending in frequency steps of 175 Hz for each second of the delay and ending in a 1 kHz tone at the same moment that the feeder released the pellet reward. Thus, whichever side the rats chose, they heard at least two tones; the first tone indicating exit from the choice point and entry into the feeder zone, and the second tone coincident with reward delivery (also indicating that the waiting period had elapsed—one second on the non-delay side, and a variable time on the delay side). In addition, each feeder (Med-Associates, St. Albans VT, USA)

made an audible click during delivery of each pellet. Therefore, the task provided no cues to the rats before they made their choice, but it did signal the reward delay as soon as they exited the choice point boundary.

The structure of the delay-discounting task encouraged animals to “titrate” the adjusting delay to their preferred delay—the delay at which waiting for a three-pellet reward was equal in value to an immediate one-pellet reward [112]. Behavioral results, including titration behavior, are described in detail in Chapter 3; Behavioral Results.

### Surgery

In order to investigate the roles of the orbitofrontal cortex and the ventral striatum on the spatial adjusting delay-discounting task, rats were implanted with electrode arrays to measure neural correlates of decision-making behavior in these two structures. After the 30-day behavior sequence, rats were chronically implanted with a “hyperdrive” (Kopf, Tujunga, CA, USA) consisting of 12 tetrodes and 2 reference electrodes that could be individually lowered into the brain across days until reaching the desired depth.

Prior to surgery, rats were anesthetized with Nembutal (sodium pentobarbital, 50mg/kg, Abbott Laboratories, Chicago, IL), their skull was shaved, and they were placed in a standard stereotactic apparatus (Kopf). Once they were secured on the stereotax, a “nose cone” was made around the rats’ nose and isofluourane gas (0.5-2 % isofluourane in oxygen) was administered throughout the surgery to maintain a stable level of anesthesia. Rats received a prophylactic injection of the antibiotic Dualcillin (0.2cc; Phoenix Pharmaceutical, St. Joseph, MI) intramuscularly in the hindlimb, and the analgesic ketoprofen (2.5mg/kg) subcutaneously. The scalp was disinfected with iodine solution and a scalpel was used to make an incision to expose the skull.

Two craniotomies were made, one for OFC, and one for vStr. A custom-made, manually drivable auxiliary electrode was placed above the hippocampus (AP -3.8, ML +2.5, relative to

bregma) and lowered over days into the area near the hippocampal fissure (where theta activity is strongest). A wire wrapped around a screw that was placed above parietal cortex was used as ground. The hyperdrive was gently lowered into place above the craniotomies so that the tetrodes (which extended out 1mm beyond their cannulae) just touched the brain surface. Small anchor screws were placed around the skull to provide points of contact for the dental acrylic, which was built up in a mound around the hyperdrive to secure it in place. After the completion of surgery, rats were given saline and Baytril (an antibiotic), subcutaneously. The tetrodes were lowered into the superficial cortex, and the rat was placed in a temperature-controlled incubator for several hours to recover. Rats were given children's Tylenol orally once they were conscious, and they were returned to the vivarium once they were sternal and able to eat and drink. Rats received Baytril and saline for five days post-surgery.

All hyperdrives contained two separate bundles of tetrodes, one bundle targeting lateral orbitofrontal cortex (OFC, coordinates: AP +3.5, ML +2.5 mm relative to bregma), and one bundle targeting ventral striatum (vStr, coordinates: AP +1.8, ML +2.0 mm relative to bregma). Rats had either 6 OFC tetrodes and 6 vStr tetrodes (n = 4 rats), or 4 OFC tetrodes and 8 vStr tetrodes (n = 2 rats). Reference electrodes for vStr were placed in corpus callosum, and for OFC they were placed in corpus callosum or a quiet region of cortex above OFC. Neurophysiological data were collected from 164 individual recording sessions. Although all six rats ran 30 days of the delay-discounting task after surgery, we did not begin recording after surgery until the rats were running close to 100 laps plugged in and the tetrodes were close to their final targets. The distribution of recording sessions is as follows: R206 = 24 sessions, R214 = 24 sessions, R224 = 29 sessions, R226 = 28 sessions, R235 = 30 sessions, R244 = 29 sessions.

### Electrophysiological recording

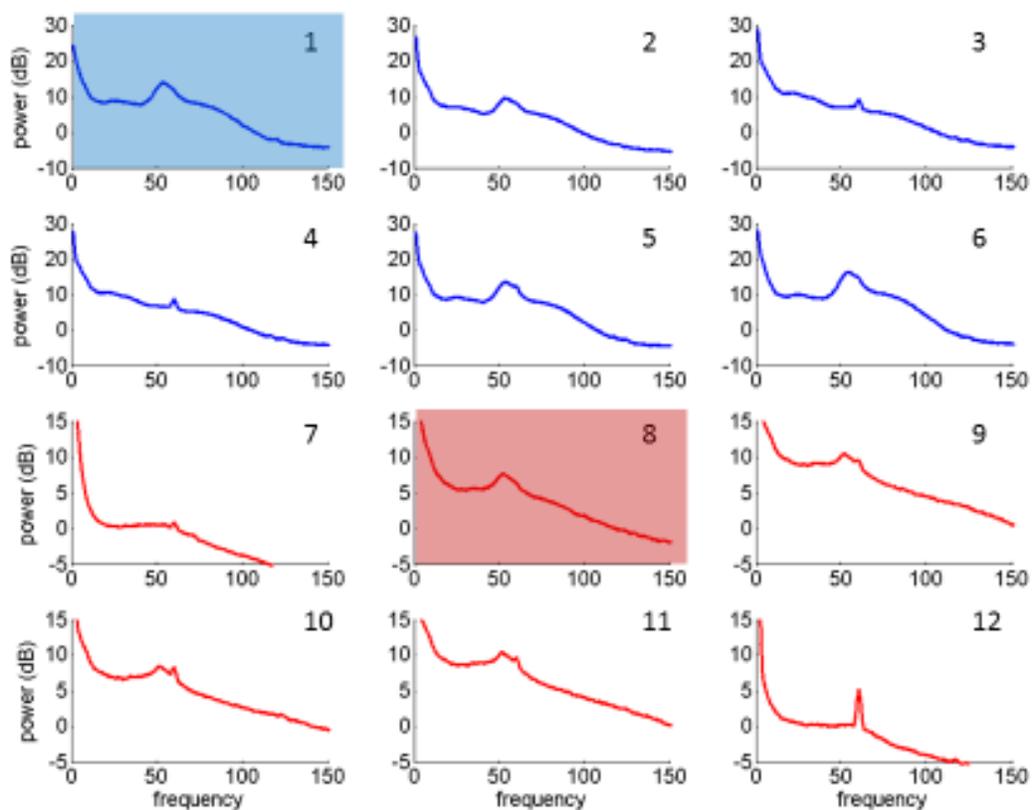
After recovering from surgery, rats once again performed maze training with one or the other side blocked off. This allowed them to acclimate to the weight of the implant and to ensure

that they could run 100 laps within the one hour session time limit. During this re-training period, which lasted between one to two weeks, tetrodes were advanced daily toward their eventual target depths: OFC (~3-3.5mm below brain surface) and vStr (~6.5-8mm below brain surface) [145]. See Figure 2-3 for final tetrode positions. Once rats reliably ran 100 laps to the left and right sides, they began the 30-day sequence on the delay-discounting task while plugged in. Each session involved a 5 minute period of recording before and after the task, while the rat sat on a ceramic pot (in the same room). After each recording session, tetrodes were either kept in place, or advanced in small increments (40-80 microns per day) to maximize ensemble size. During recording sessions, the position of the rat was tracked by an overhead camera (sampled at 30Hz) using LEDs on the recording headstage. All position data were time-stamped by the Cheetah data acquisition system (Neuralynx, Bozeman, MT). Code for running the task was custom written in MATLAB.

Single unit and local field potential (LFP) data were recorded using a 64-channel Cheetah recording system using standard techniques. For spike trains, when signals exceeded a manually determined voltage threshold (usually 50 $\mu$ V for OFC tetrodes, 40 $\mu$ V for vStr tetrodes), waveforms were sampled at 32 kHz for 1ms (filtered at 600-6,000 Hz). For LFP data, signals were continuously sampled (Continuously Sampled Channels, or CSCs) from one of the four wires on each tetrode (usually channel 1, unless the channel was noisy) at a rate of either 1990.4 Hz or 1995 Hz (the latter corresponding to a software update in the Cheetah system) and band-pass filtered between 1 and 475 Hz. A notch filter at 60 Hz to reduce electrical line noise was applied. Electrophysiological data were recorded to disk for offline analysis. Candidate clusters of putative single cells were estimated automatically using Klustakwik 1.7 [146] (K. Harris, available at <http://klustakwik.team.github.io/klustakwik>). Final categorizations of single units were identified manually using the MClust 3.5 spike sorting software suite (A.D. Redish, software available at <http://redishlab.neuroscience.umn.edu/mclust/MClust.html>). Only cells with more than 100 spikes were included in analyses.

Only one CSC each for vStr and OFC was selected for analysis from each session. Selection was determined prior to any LFP analysis. CSCs were chosen manually based on the power spectral density plots for all CSCs for each session. For the hippocampal electrode, if there was a discernable peak at theta, then it was marked for analysis. Out of 140 sessions used for LFP analyses, 89 sessions had an acceptable theta peak in the hippocampal electrode. For the OFC, CSCs with the greatest “bump” at gamma50 and the least noise (such as noise at 60 Hz or its harmonics) were selected for analysis. For vStr, the same criteria were used, but also taking into account the canonical shape of the shoulder at gamma80 (as can be seen in Figure 4-1 B). Figure 2-2 shows example power spectra from all tetrodes from a representative session.

A



B

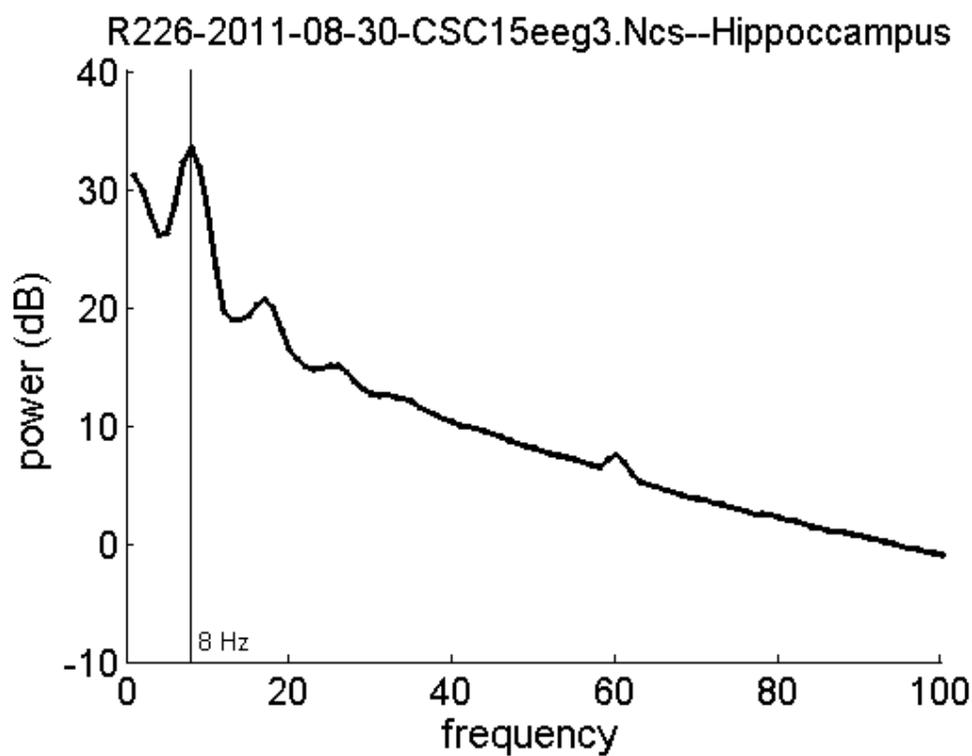


Figure 2-2. **PSDs from an example session.** (A) Power spectral density plots for the 12 tetrodes in vStr (blue: 1-6) and OFC (red: 7-12) from session R226-2011-08-19. The CSC channels that were chosen for analysis are highlighted in color. (B) Power spectral density plot for the auxiliary electrode in hippocampus for the same session. Note the peak in theta power at 8Hz. This CSC was classified as having theta, and was used for analysis.

Beyond the computational necessity, this approach is justified on several counts. When analyzing data for single sessions, during the coding process, similar results were obtained using different CSCs within a given brain structure, so long as the CSC was not strikingly noisy (i.e. a “bad” tetrode). The selection criteria were based on choosing CSCs that had noticeable power in the gamm50 band for OFC, and in the gamma50 and gamma80 band for vStr. Oscillations at these frequencies are signature features of the LFP signal from these brain structures [119, 128, 129, 147-151]. Most importantly, the results from my LFP analyses, where they can be compared directly (such as PSDs, self-coherence, spectrograms, and gamma power dynamics), are in agreement with other reports in the literature, both from our own group [119, 152], and from other labs [128, 129, 148, 149]. This correspondence between results lends confidence to the technique used here. While some reports have emphasized the differences in LFP signals in the vStr between different tetrode locations [149], the broad and reliable trend of a dorsolateral to ventromedial gradient [147] in signature frequencies of the power spectra is a much stronger effect.

After completion of the electrophysiological recording sequence, rats were sacrificed and their brains were sliced and stained using standard histological techniques (see Histology, below). For single-unit and local field potential analyses, R214 was excluded (see below for explanation). This yielded 140 sessions from five rats for single-unit (947 vStr cells, 1754 OFC cells) and local field potential (140 CSCs each for OFC and vStr, 89 for hippocampus) analyses. For Bayesian decoding analyses, we only included sessions with at least five cells each in vStr and OFC; 85 sessions (177 vStr cells, 681 OFC cells).

### Histology

Following the completion of neurophysiological recording, the final locations for each tetrode were marked, or “gliosed,” by applying a small current injection through two of each of the four tetrode channels using a stimulus isolator—10 $\mu$ A DC current, 10 seconds for each channel. Passing a small amount of current into the brain causes reactive gliosis, which shows up readily under a cresyl violet stain. References were not marked, so as not to confuse them with the recording tetrodes. Any tetrodes that did not show biological signals during the experiment were not marked. These tetrodes were easy to identify as they were characterized by large noise artifacts during recording, a lack of cells, and noise on the LFP channels, and this was almost always present from the beginning of recording. Only a handful of tetrodes met these criteria, and they were not used for subsequent single unit analyses (by virtue of them not having well-isolated cells) or local field potential analyses.

At least two days after the gliosis procedure, rats were deeply anesthetized with sodium pentobarbital and perfused transcardially with a solution of 10% formalin in saline. The brains were removed, placed in formalin overnight, and stored in a solution of 30% sucrose in formalin until they were sliced. Brains were blocked into a section that included OFC and striatum (and adjacent areas anteriorly and posteriorly) and sliced in the coronal plane at 40 micron thickness on a freezing microtome. Individual sections were laid onto cover slips and underwent a cresyl violet stain to visualize electrode tracks and gliosis marks.

Stained slides were examined under a standard optical microscope and tetrode locations were manually recorded. Marks were made at those locations on corresponding anatomical images (coronal sections) from the atlas of Paxinos & Watson [145]. All OFC tetrodes fell within the lateral orbital and ventral orbital cortices (Figure 2-3a). OFC histology for R224 was unavailable due to a technical error during brain slicing. However, analysis results for R224 matched those of the other rats. Tetrode endpoints for R214 entered into piriform cortex. Therefore, R214's electrophysiological data was not used for analysis, and his tetrode positions are not shown. R214's behavioral data from his recording sessions was used for behavioral analyses. All tetrodes (with

one exception) from the striatal bundles were located within vStr or the ventral caudate/putamen. Most striatal tetrodes fell within the nucleus accumbens core, with a few in the more lateral aspect of nucleus accumbens shell (Figure 2-3b).

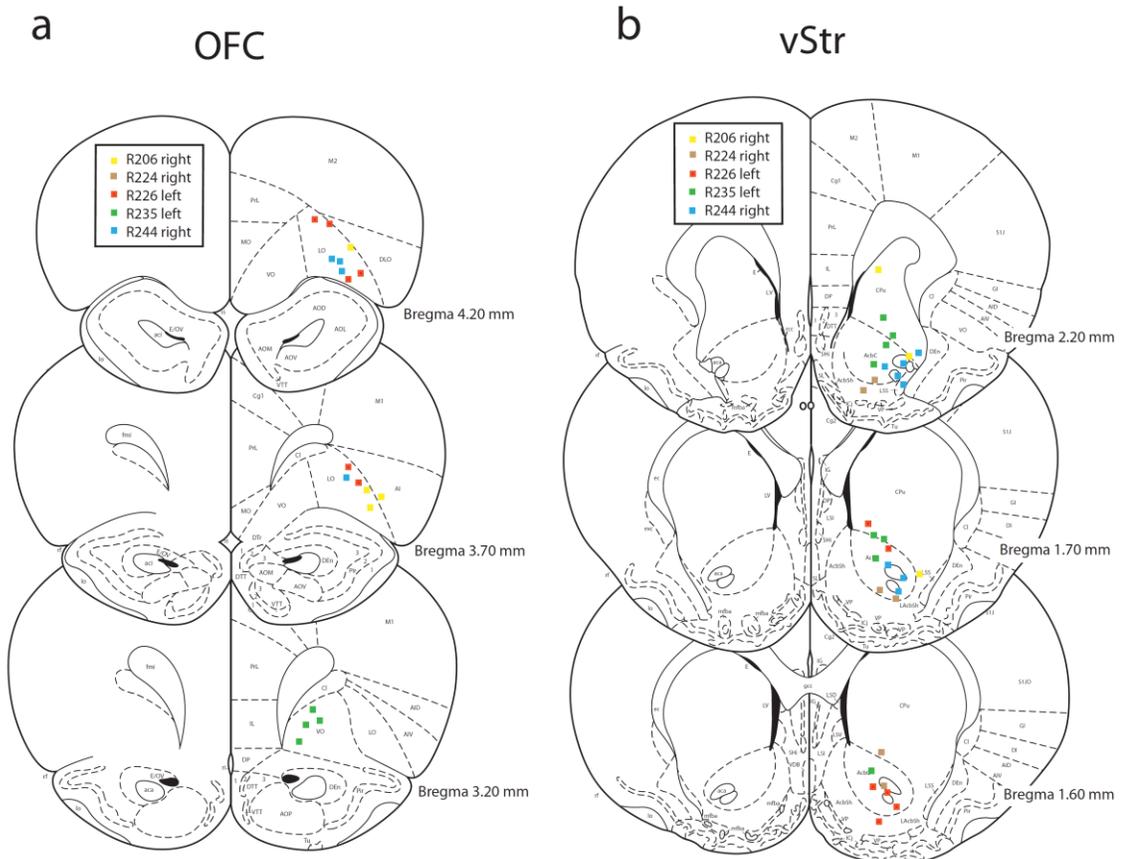


Figure 2-3. **Histological determination of recording sites.** (a) Final positions of OFC tetrodes as determined by visual inspection of tetrodes tracks and gliosis marks. Anatomical drawings are from Paxinos & Watson, 1998 [145]. OFC tetrodes fell within the lateral orbital and ventral orbital cortices. (b) Final positions of vStr tetrodes. Tetrodes were located within vStr or ventral caudate/putamen. Figure image from [144], reprinted with permission.

## Data Analysis

### Reward-sensitivity

To determine the reward sensitivity of a neuron, we calculated its mean firing rate in a window from 0 to 4s after feeder trigger events. This time window encompassed the approximate time of reward receipt and consumption. For each neuron, we determined reward-sensitivity for the delay-side feeder, the non-delay side feeder, and for both feeders taken together. In each case, we calculated a bootstrap distribution by determining the firing rate of the same neuron at random times during the session, in 4 second windows, for as many laps as the rat made to that feeder(s). See Figure 2-4 for an example neuron that illustrates this method. We created a bootstrap distribution (500 iterations) by running this same algorithm using random time windows instead of the feeder times. For example, if the rat visited the non-delay feeder 48 times during a session, this distribution tells us what to expect from averaging the firing rate for the same neuron over 48 random times within the session. Neurons were considered *reward-responsive* if the neuron's mean firing rate during reward receipt was significantly different from the bootstrapped distribution (z-test,  $p < 0.05$ ). To visualize the reward-related response of the entire population of recorded cells, heat maps were generated for vStr and OFC, showing the firing rate response of each cell to the large reward feeder, the small reward feeder, or both feeders taken together. Cells are sorted according to their z-scored firing rate response (calculated as described above) to both feeders (Figure 2-5).

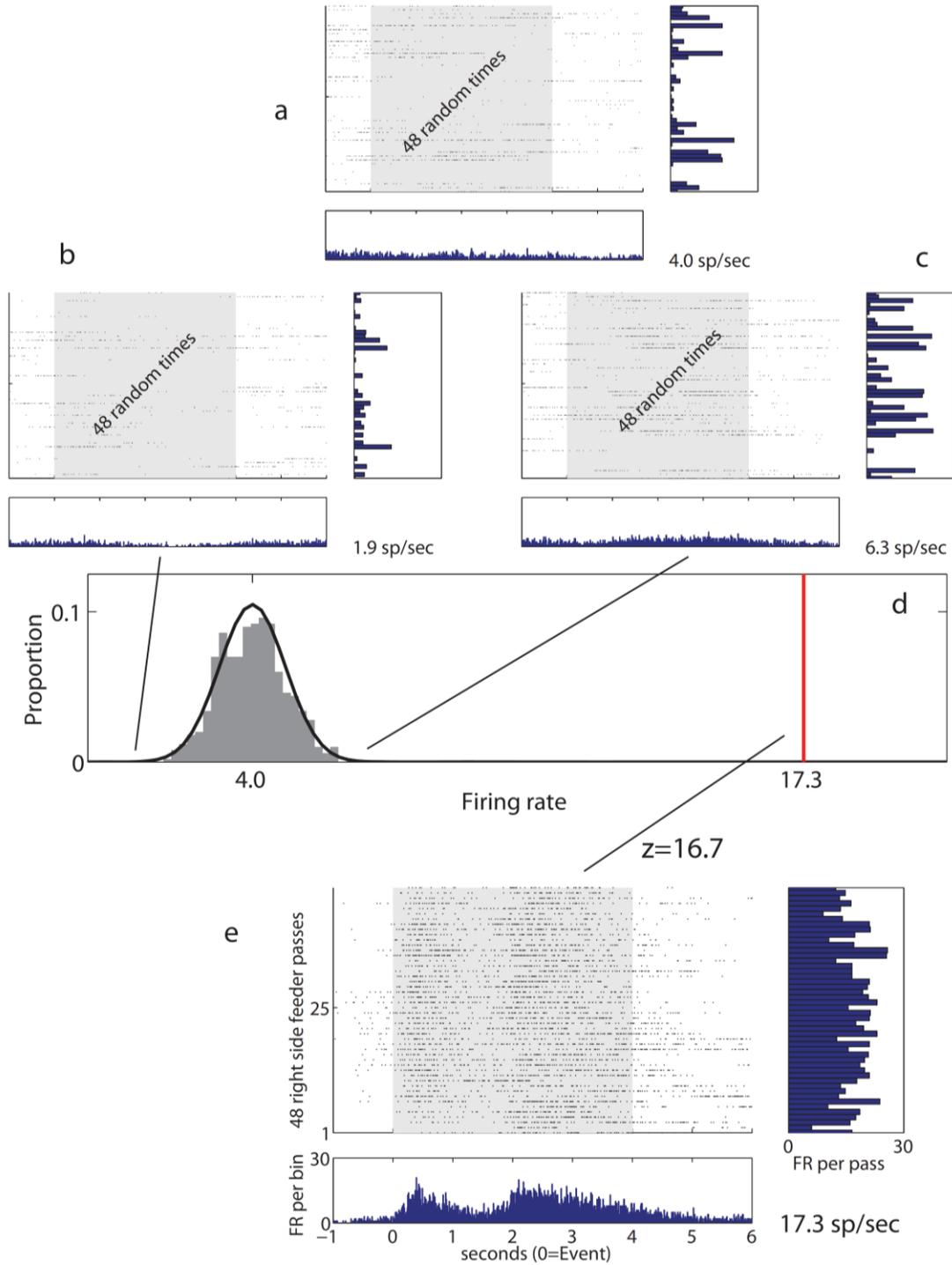


Figure 2-4. **Schematic diagram of procedure for calculating z-scores.** This data was taken from session R235-2012-01-08. For each neuron, the responses to reward during the response window (0-4 s after feeder fire) at either of the feeders are not a single number, but a distribution. In this

session, the rat visited the right side feeder 48 times, and the spike raster and distribution of firing rates is shown in **(e)**. We took the average of that distribution to get an average firing rate for that cell to the right feeder (17.3 spikes/s). We next calculated a comparison distribution of firing rates for the same cell during the same window (0-4 s after feeder fire), but taken at random times during the session. Importantly, we used the same number of samples for the comparison distribution as the rat actually experienced at the right feeder; namely, 48. We then took the average of that comparison distribution to get an average firing rate. Any one sample of 48 random times within the session could yield uncommonly low or high average values compared to the overall baseline firing rate for that cell (which was 4.02Hz)—exemplified in panels **(b)** and **(c)**, respectively. Therefore, we repeated the process 500 times to get a distribution—the bootstrap distribution (shown in panel **(d)**, although this particular example only contains 100 iterations)—which gives us a valid estimate of the distribution of firing rates from that cell when it is sampled 48 times at random. A sample taken from the middle of the bootstrap distribution, shown in panel **(a)**, lines up very well with the true baseline firing rate of that cell (4.02 Hz). The key feature of this approach is that we performed the same operation to generate the bootstrap distribution (and its attendant mean and standard deviation) as we did to find the reward response for the given cell. Separate bootstrap distributions were created for the left and right feeders, as well as for both feeders taken together. For this particular neuron, it had an average firing rate for the right feeder of 17.3 spikes/s, giving it a z-score of 16.7, compared to the bootstrap distribution—red line in **(d)**. Figure image from [144], reprinted with permission.

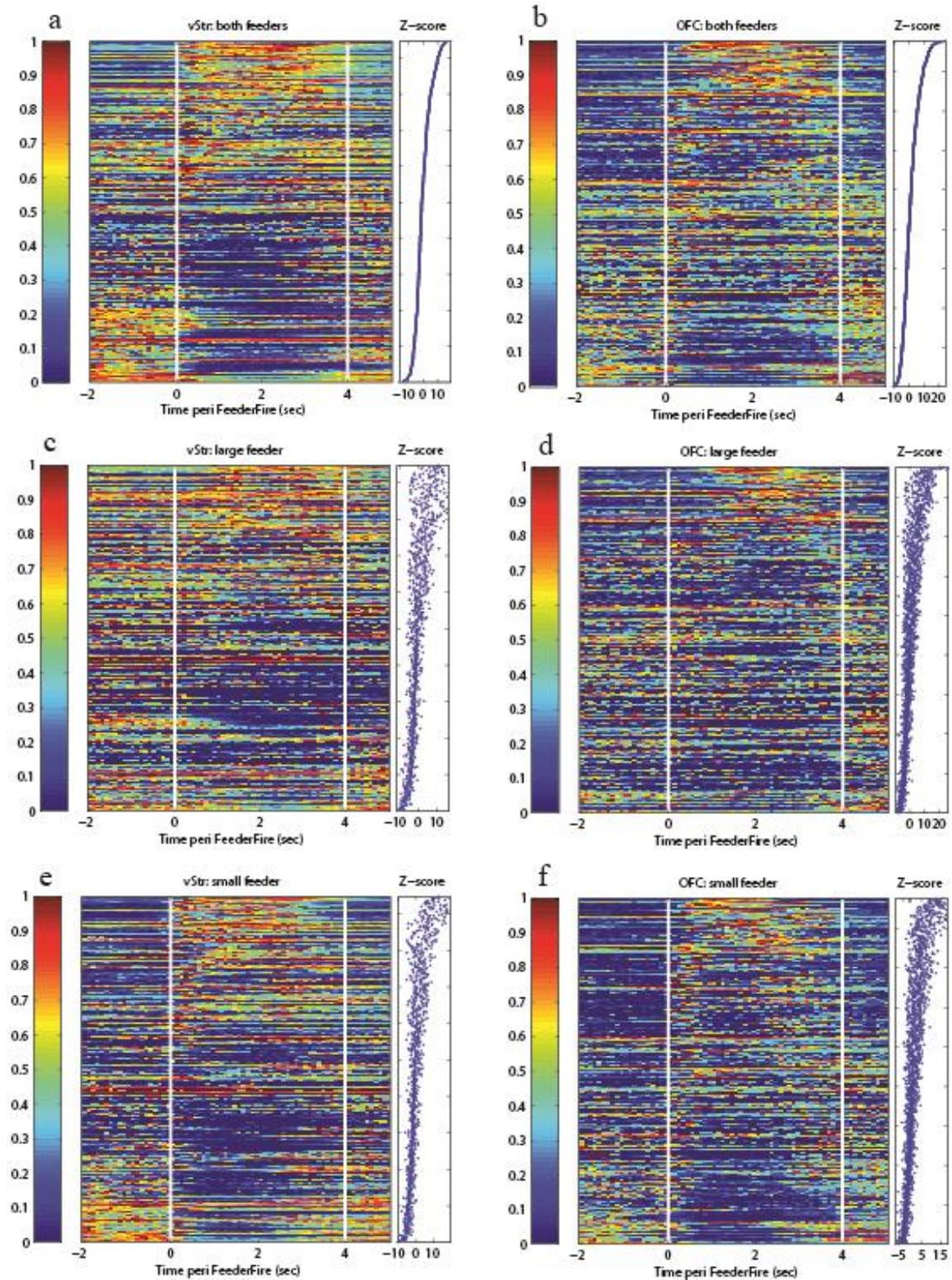


Figure 2-5. **Heat/line plot showing raw firing rate data around reward receipt.** The heatmap in the middle of each plot shows the firing rate response of all neurons within a brain structure during

the time window of -2 to 5 sec around reward delivery. The top row (**a,b**) shows the response to both feeders combined. The middle row (**c,d**) shows the response to the large feeder only. The bottom row (**e,f**) shows the response to the small feeder only. These firing rate responses are normalized to the cell's maximum firing rate (range of 0 to 1) and they are ordered vertically by the z-score that defined each cell's reward response to both feeders (high z-scores at top, descending). The reward response window is marked by white lines. At right in each plot is the z-score for each neuron to the feeder(s) specific to that plot. The z-scores are aligned to those from the top panel (sorted z-scores for both feeders), and therefore do not appear as a uniform line, as on the top panel. Figure image from [144], reprinted with permission.

### Identification of vicarious trial and error

Headstage tracking with LEDs allowed precise measurement of the head position of the rat during passes through the choice point. Position samples starting from the midpoint of the central stem of the maze and ending at the invisible line demarcating entry into the feeder zone defined the choice point window (Figure 2-1, dashed lines through the central stem). The coordinates defining the choice point zone were identical from day to day. Only position data from the choice point zone were used to categorize laps as VTE or non-VTE. Episodes of VTE were identified by calculating the *curvature* of the trajectory through the choice point, measured as the tortuosity of the trajectory [153]. In order to calculate the curvature at each moment in time, we started from the  $\langle x, y \rangle$  sequence of position samples detected from the headstage via the camera in the ceiling and Neuralynx's position tracking software. From this sequence, we calculated the velocity  $\langle dx, dy \rangle$  by applying the Janabi-Sharifi algorithm to the position sequence [154]. From the velocity sequence, we calculated the acceleration  $\langle ddx, ddy \rangle$  by applying the Janabi-Sharifi algorithm to the velocity sequence [154]. We then calculated curvature as defined by the formula for tortuosity [153]:

$$(dx * ddy + dy * ddx)/(dx^2 + dy^2)^{1.5} \quad \text{Equation 2-1}$$

A sequence of adjacent tracking points with consecutive curvature values greater than 2 defined a re-orientation event. From visual inspection, choice point passes with curvature values greater than 2 matched well with subjective determination of VTE events.

The start of the re-orientation event was found by taking the position that started 200ms before the first position sample with curvature greater than 2. This “TurnAround” point matched well with the qualitative judgment of the moment of re-orientation (see Figure 2-6, c & d). When multiple, discrete re-orientation events occurred within some individual choice point passes, only the first one was analyzed. Any choice point pass with a maximum curvature value greater than 2 was defined as a VTE lap. For laps with curvature values below this threshold, the point of alignment was taken as the halfway point through the choice point trajectory (called the “MidPoint”) (Figure 2-6, a & b). Under this classification system, using all behavioral data, 2,099 out of 16,000 laps (13%) were classified as VTE laps (n = 164 sessions). See Figure 3-3 for the distribution of VTE and non-VTE laps.

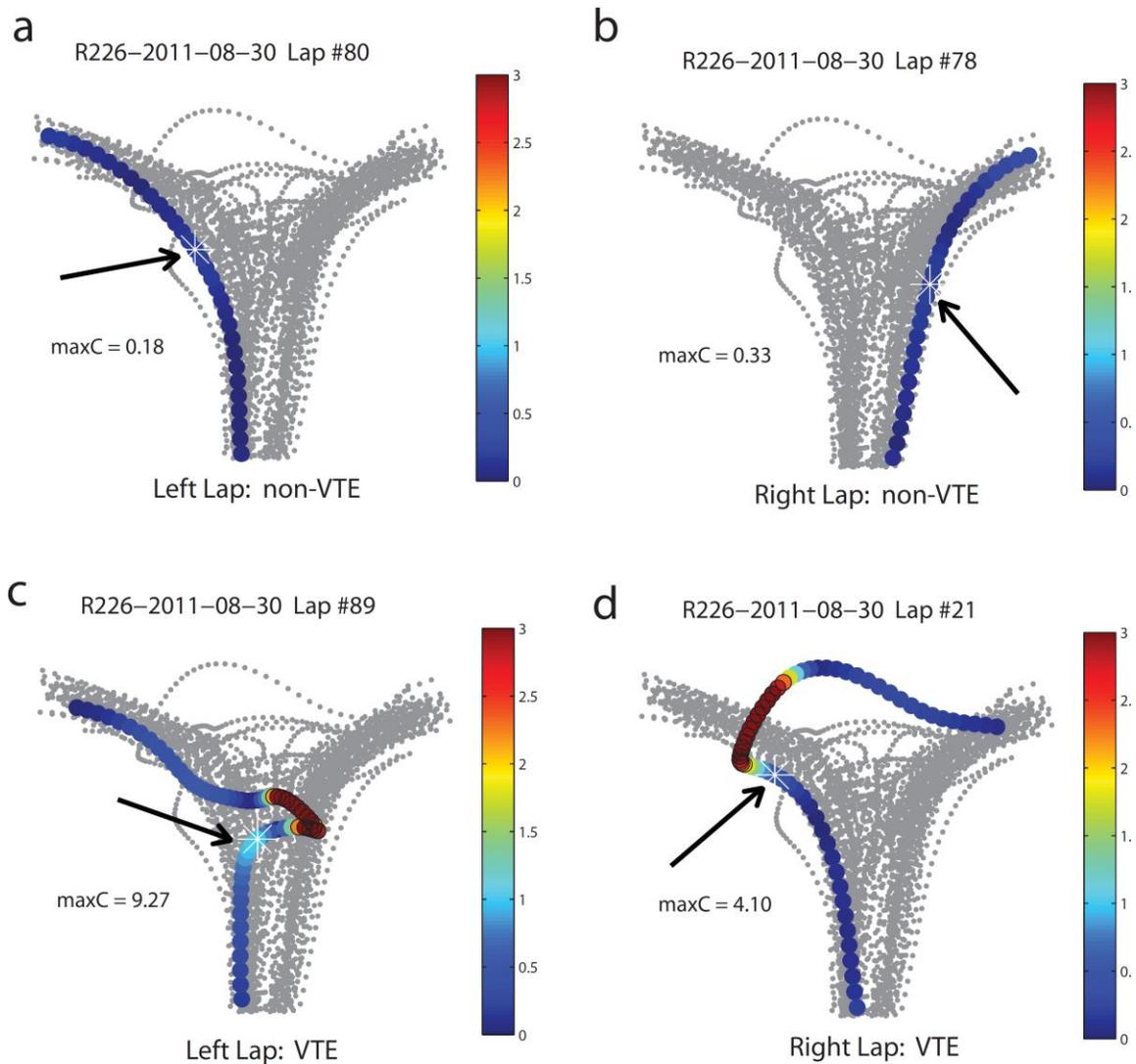


Figure 2-6. **Example non-VTE and VTE events.** Each of the four plots shows individual behavioral examples taken from one recording session for R226. Position data from all laps are plotted in the background in grey. Over the course of a session, the rat traced out two broad curving trajectories through the choice point (one to the left, one to the right). Individual laps of interest are plotted in color. Warmer colors correspond to higher curvature values. The maximum curvature value (maxC) for the highlighted laps are shown at left in each plot. Laps were classified as VTE laps when the maximum curvature was greater than 2. **(a)** and **(b)** show non-VTE examples of left and right passes through the choice point, respectively. The MidPoint for each lap, defined as the halfway point of the trajectory, are indicated by arrows pointing to the white asterisks. **(c)** and **(d)** show VTE examples of left and right passes through the choice point, respectively. Arrows point

to the white asterisks, which indicate the start of the “TurnAround” for each VTE event. Figure image from [144], reprinted with permission.

### Ensemble decoding

On the spatial delay-discounting task, reward receipt only occurred at the feeder sites on the maze. Thus, reward-related activity only occurred at those same feeder sites. In order to measure representations of those reward-receipt locations, we used a spatial decoding approach, in a manner similar to that used previously [48, 120]. All decoding analyses were performed using a one-step spatial Bayesian decoding algorithm [155] with a time step of 250ms non-overlapping windows and a uniform spatial prior. We used a “leave-one-out” procedure so that the decoding was done lap by lap. For each lap, the training set for the decoder (derived from the tuning curves) was calculated from all activity excluding the current lap (i.e. using all other laps). In this way, the training set for the decoder did not include the test set being decoded.

This algorithm is an implementation of Bayes’ rule:

$$p(x|s) = p(s|x) * p(x)/p(s) \quad \text{Equation 2-2}$$

$P(x|s)$  is the measure of interest, called the “posterior,” or “posterior probability.” It is the estimated position of the rat, although, as we will see, it can also tell us how similar the spiking activity at any given time is to the average spiking activity at the reward sites.  $P(s|x)$  is the tuning curve of the neuron.  $P(x)$  is the likelihood of the rat being at position  $x$ , independent of the spike count, and is called the “prior,” or “prior probability.” We used a uniform, or “historyless,” spatial prior, meaning that the total probability (equal to 1) was evenly divided across all spatial bins.  $P(s)$  is a probability distribution that tells us the likelihood of a given neuron emitting  $n$  spikes, assuming a Poisson spiking model. The decoding algorithm averages the probabilities across all neurons in the ensemble to give a single posterior value for each spatial position  $x$ .

It is important to note that the use of this algorithm does not imply (nor does it depend on) orbitofrontal or ventral striatal cells having any spatial firing correlates. Reward-related correlates will “drag” the spatial decoding to the reward sites [48, 120].

In order to measure the reward-site representation, we first identified what positions the animal sampled during the four seconds (0-4 s) after reward-receipt. For each session, we defined the feeder sites as those spatial bins where there was greater than zero occupancy during the reward response window (0-4 s after feeder fire). We then calculated the proportion of the posterior probability allocated to those feeder sites. Thus, this decoding method measured the probability that neural ensemble activity during the given time step decoded to the particular spatial locations occupied by the feeders on the maze. Again, this does not imply or require that vStr or OFC cells are spatial in nature. Rather, cells in vStr and OFC respond to specific events (i.e. reward receipt); they show reward tuning, and these events happen at specific spatial locations (i.e. at the feeders). Because reward cells tend to fire at the feeder sites, an algorithm initially designed to make predictions about the rat’s position in space [155-157] can also tell us the probability that the rat’s neural activity at that moment is representing the reward sites [48, 120].

For each time step, we defined  $pFeeders$  as the mean of the probabilities of all the bins that constituted the feeder sites. All analyses divided the positional tracking data into a grid of 32 x 32 spatial bins (this includes the maze and adjacent space within the camera’s field of view). Because the absolute values of the probabilities obtained depended on the number of spatial bins used (e.g. 16 x 16 versus 32 x 32), we included a normalization factor that kept  $pFeeders$  values constant, independent of the number of spatial bins used. To normalize for bin number (and thus bin size), we multiplied the  $pFeeders$  values from each session by the number of bins in which the rat spent any time in that session (occupancy greater than zero). This normalization procedure meant that the expected  $pFeeders$  values from a uniform posterior would be 1. Values greater than 1 indicate higher than chance levels of decoding to the feeders; values less than 1 indicate lower than chance levels of decoding to the feeders. Comparisons shown in Figure 3-8 are made to  $pFeeders$  values

derived from shuffled data (in which the spike order of each individual cell was randomized, maintaining the first order spiking dynamics of the cell) [158].

### Power spectral density

Power spectral density (PSD) plots were calculated for each session with a Welch spectrum estimator; *spectrum.welch* from the MATLAB Signal Processing Toolbox, with a single Hamming window of 512 in non-overlapping windows. Unless otherwise noted, data were kept at their original sampling rate (~1990 Hz). Whole-session PSD vectors (n = 140) were averaged and plotted with the mean  $\pm$  SEM for Figure 3-1.

### Self-coherence

As reported in [159], fundamental frequencies in the power spectrum and their interrelationships can be visualized by plotting the cross-correlation of the power spectrum. Here, spectrograms for each session were calculated using the built-in MATLAB function *spectrogram*, with a Hamming window of 1024 (~0.5 sec window) and 50% overlap, from 0 to 300 Hz at 1 Hz resolution. The correlation (*corrcoef* in MATLAB) across frequencies was calculated using all time segments (of length roughly 4x the number of seconds in the session, because the spectrogram window equals 0.5 sec with 50% overlap) to generate a 300 (Hz) x 300 (Hz) “self-coherence” matrix of correlation coefficients. Data were averaged across sessions to generate the self-coherence plots in Figure 4-3. Values along the diagonal are, by definition, equal to 1. Warmer colors off the diagonal indicate high correlations coefficients; i.e. a correlation in power at those frequencies. Cooler colors indicate low correlation coefficients, meaning little to no correlation in power at those frequencies. Values below zero (dark blue) indicate an anti-correlation in power at those frequencies. The two halves of the plot (upper and lower triangles, bounded by the identity line) are symmetrical by definition, but both halves are included for the sake of visualization.

### Gamma power estimation

In order to measure the temporal dynamics of gamma power in OFC and vStr, raw CSC time series were converted to units of microvolts and bandpass filtered. The frequency range of 45-55 Hz was used for gamma50 and 70-85 Hz for gamma80 in order to allow for a direct comparison with previous work on gamma power dynamics [119, 149]<sup>9</sup>. After filtering, the Hilbert transform was taken to obtain estimates of instantaneous gamma power at each time point. This new time series of gamma power was restricted to specific times around the event of interest (TurnAround, MidPoint, or Feeder fire events) and concatenated across sessions. Data were averaged, and plotted as the mean  $\pm$  SEM for the peri-event time histograms (PETHs) (Figures 4-5, 4-6).

### Coherence

Multi-taper coherence between paired LFP traces was calculated using the function *cohgramc* from the Chronux toolbox (freely available at [chronux.org](http://chronux.org)) [160]. The multi-taper approach seeks to optimize the tradeoff between narrowband and broadband bias that is inherent in spectral estimation. See [161] for a discussion of the advantages of this technique. In practice, tapers reduce the influence of distant frequencies at the expense of smoothing the frequency spectrum over nearby frequencies. The example spectrograms and the coherogram in Figure 3-4 were calculated with parameters identical to that used in [119] Figure 4B. Smoothing is evident, especially at the lowest frequencies (~1-20 Hz). Importantly however, the takeaway from that figure—namely, the approximate duration of gamma bouts and their alternating pattern—have been demonstrated using other techniques [148, 149], and is also evident in the raw CSC illustrated in Figure 4A from that paper [119].

---

<sup>9</sup> For the coherence analyses, gamma80 was defined as 70-100 Hz, as this best matched the shape of the average coherence spectra in this frequency range (Figure 4-7 F).

The calculation of coherence used for the session-wide averages (Figure 4-7) and for coherence PETHs (Figures 4-8 & 4-9) used more conservative parameters in order to obtain better spectral resolution. Specifically, cross-spectral coherence was calculated on zero-padded data in a moving 1 second window with 20% overlap, from 0-300 Hz, with a bandwidth of 5 Hz. For average coherence spectra, data were averaged across all time bins to give a session-level average, and the matrices from all sessions from each rat (Figure 4-7 A-E) or all rats taken together (Figure 4-7 F) were averaged to obtain the mean coherence spectra  $\pm$  SEM. For the PETHs in Figures 4-8 and 4-9, coherence data were averaged within frequency bands (45-55 Hz or 70-100 Hz) to obtain gamma50 coherence and gamma80 coherence vectors. These data were restricted to the window of interest and concatenated, then averaged to obtain the mean coherence  $\pm$  SEM, aligned to their respective event times (TurnAround, MidPoint, Feeder Fire), for each condition (VTE versus non-VTE).

### Granger causality

Granger causality analyses were performed by customizing code based on core functions from the Multivariate Granger Causality Analysis Toolbox (MVGC) [162], available for download at [users.sussex.ac.uk/~lionelb/MVGC](http://users.sussex.ac.uk/~lionelb/MVGC). This toolbox is a more recent version of a previous release, and both toolboxes have been used extensively in the neuroscience literature in recent years [163]. As described by Barnett and Seth, Granger causality (“GC” or “G-causality” for short) is based on the ideas of predictability and precedence [164]. If signal A precedes signal B, and if signal A predicts signal B above and beyond what past values of signal B predict about its present state, then A can be said to “Granger cause” signal B. G-causality was operationalized by the economist Clive Granger (Granger, 1969), who used vector autoregressive (VAR) models to fit time-series data. VAR modeling entails finding a linear combination of optimal weights that minimizes the prediction error of subsequent values. This approach was extended into the spectral domain by John Geweke [165, 166]. Both “time-domain” and “spectral-domain” GC analyses are used in this

manuscript. Spectral-domain GC is a decomposition of time-domain GC values into their spectral components.

For each session, two time series from paired OFC and vStr electrodes were used (selected manually, before any LFP analyses were conducted). A principle condition of VAR modeling is that the data is “stationary.” Chiefly, this means that there are few autoregressive components in the time series. Figure 2-7 illustrates the process used here to reduce autocorrelation to an acceptable level. Panels a & b show raw traces from OFC and vStr, along with plots of the autocorrelation coefficients for each trace, from zero to 100 time steps. Although the sampling rate in this example is 1990 Hz, the data used in these analyses were downsampled by a factor of 4 (to ~500 Hz), to avoid overfitting. Panels a & b show strong autocorrelations (which should be roughly less than  $\pm 2$  divided by the square root of the number of data points, indicated by the dotted red line). In order to reduce the autocorrelation, each data point in the raw trace is subtracted from the next. This is called “differencing,” and is a standard technique for GC analysis [167]. First-order differencing dramatically reduces the autocorrelation (Figure 2-7 C), but not to the desired level (red dotted lines).

To get the autocorrelation within acceptable bounds, the differenced data was “pre-whitened” by applying an autoregressive integrated moving average (ARIMA) model, which explicitly removes autoregressive components in the time series [168]. We used a model order of 25 (25 autoregressive terms) and 1 moving average term [25 1]. The ARIMA technique has been used fruitfully to analyze correlations between time series data [169], and pre-whitening in general has been argued to be an essential step for time series analysis [170]. What is left after the ARIMA step are called “residuals,” which can be seen to have autocorrelation coefficients within the acceptable bounds. All GC analyses were performed on these residuals.

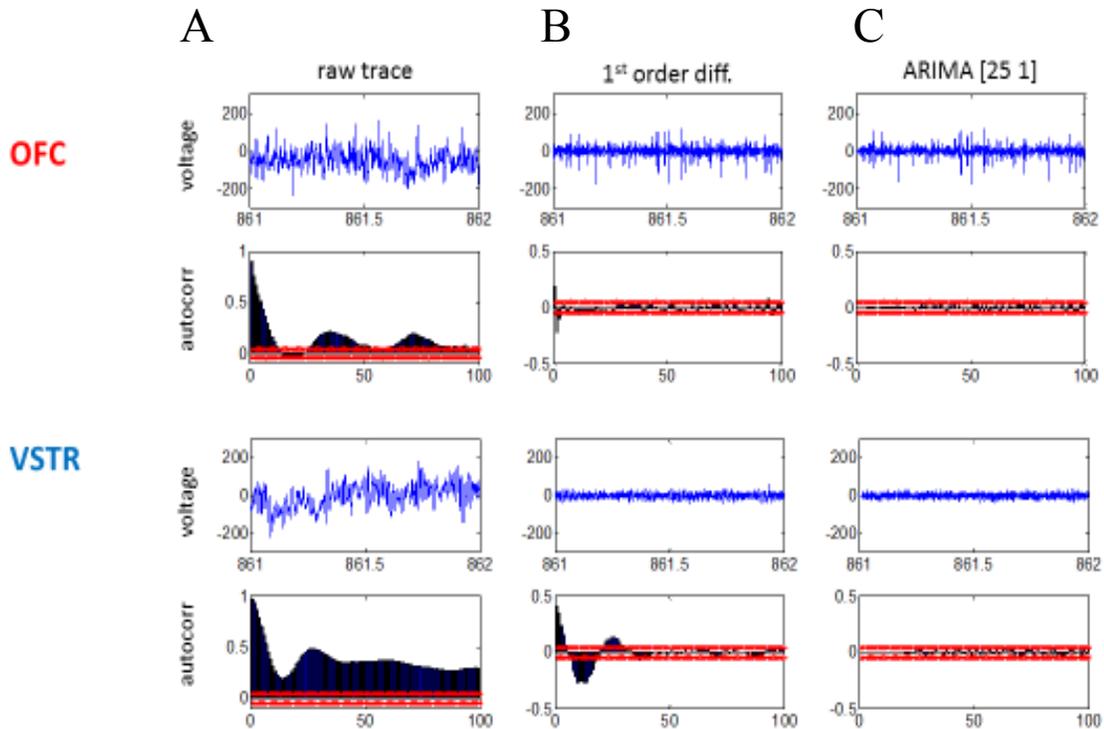


Figure 2-7. **Example of LFP data “pre-whitening” for Granger causality analysis.** Voltage traces in blue and autocorrelation values in black from an example time window on the delay-discounting task. **(A)** Raw data (blue) from OFC (first panel) and vStr (third panel). Regular oscillations are present in the sample, causing high autocorrelation values, plotted in the second and fourth panels, with lags on the x-axis and correlation coefficients on the y-axis. Red dotted lines indicate  $\pm 2/\sqrt{\text{number of data points}}$ . **(B)** First-order differencing greatly reduces the autocorrelation. **(C)** Applying an ARIMA model to the time series reduces the autocorrelation to within acceptable bounds.

A model order of 20 was used for all GC analyses, which corresponds to a time lag of 40ms (since each time step is 2ms). The model order determines how many steps the GC equations “look back” in predicting future values. This is not to say that the algorithm looks back at the 20<sup>th</sup> value, but rather that it incorporates all the data points going back from 1 to 20. Therefore, GC does not provide an explicit estimate of the lag between two neural signals, but it is constrained by the model order parameter. 40ms is long enough to include axonal and synaptic delays between vStr and OFC, and is therefore physiologically plausible. Higher model order can yield better predictions, but at

the cost of overfitting. A model order of 20 was chosen as a suitable compromise between not enough information and possible overfitting.

For Figure 4-10 A, GC values were calculated over the entire session, each session yielding one data point for “time-domain GC” with two values (OFC→vStr & vStr→OFC). Likewise, in Figure 4-10 B, each session yielded two frequency vectors for “spectral-domain GC” (OFC→vStr and vStr→OFC). Time-domain data are plotted as a scatterplot (Figure 4-10 A), with values above the identity line corresponding to a greater influence of vStr on OFC, and values below the identity line corresponding to a greater influence of OFC on vStr. Spectral data, calculated from 0 to 250 Hz (250 Hz being the Nyquist frequency for the sampling rate of 500 Hz), were averaged across sessions to create GC spectra at the “whole-session” level (Figure 4-10 B). In principle, frequency-domain GC values should add up to the time-domain GC value for any given window of analysis (in this case, the entire session). This was indeed confirmed by a check in the MVGC code.

GC data for the PETHs in Figure 4-11 were calculated in the same manner as above; namely, by analyzing the residuals. In this case, the session-long time series of residuals was divided into one second long segments and each was analyzed separately. This yielded a series of GC values which could be aligned (i.e. restricted) to behaviorally relevant events. As with the other PETHs, the restricted data was concatenated across sessions and averaged, to yield a PETH showing the mean  $\pm$  SEM for time-domain GC values in either the OFC→vStr direction, or vice versa.

#### Amplitude cross-correlation

Another, arguably more intuitive way to examine the direction of influence between neural signals is to measure their cross-correlation. This approach was formalized by Adhikari et al. [171]. Called “maximum cross-correlation”, or “max crosscorr” for short, this technique can be used to measure the lead or lag between two simultaneously recorded neural signals, telling the experimenter in effect which brain structure is the sender and which is the receiver, on average, in

a given time window [171]. Designed specifically for LFP data, the authors verified this method using simultaneous recordings from the rodent medial prefrontal cortex (mPFC) and ventral hippocampus (vHIPP), in which the anatomical connectivity is well characterized. The authors found that in most cases the vHIPP led the mPFC, as indexed by their new measure. This result was corroborated by phase-locking analysis, partial directed coherence (which is similar in principle to Granger causality), and testing on a surrogate dataset (“dummy data”), all effects consistent with the known monosynaptic connections from vHIPP to mPFC [172] and the well-established role of HIPP in influencing mPFC activity [173].

The max crosscorr technique proceeds in four steps. First, the time series data is bandpass filtered at the desired frequency range. Second, the instantaneous amplitude of the filtered LFP envelope is calculated with a Hilbert transform. Third, the cross-correlation between the two amplitude vectors is computed (*xcorr* in Matlab), yielding a series of correlation coefficients at the specified lags—usually 100ms to +100ms (but this can be set by the user). Finally, the maximum value of the cross-correlation coefficients is taken as the best estimate of the interaction between the two time series, providing both an absolute value of the cross-correlation, but more importantly, an estimate of the lag between structures, which by necessity also indicates which structure is leading the other. These steps are illustrated in Figure 2-8 with an example of simultaneous LFP traces from OFC and vStr. When many such observations are combined, they can be combined into a histogram, as in Figure 4-12, which gives the distribution of maximum cross-correlations and an estimate of the lag and directionality between structures from a collection of recordings. Significance is determined by testing the difference between the mean or median of the distribution of lags against the null hypothesis that the distribution is centered around zero.

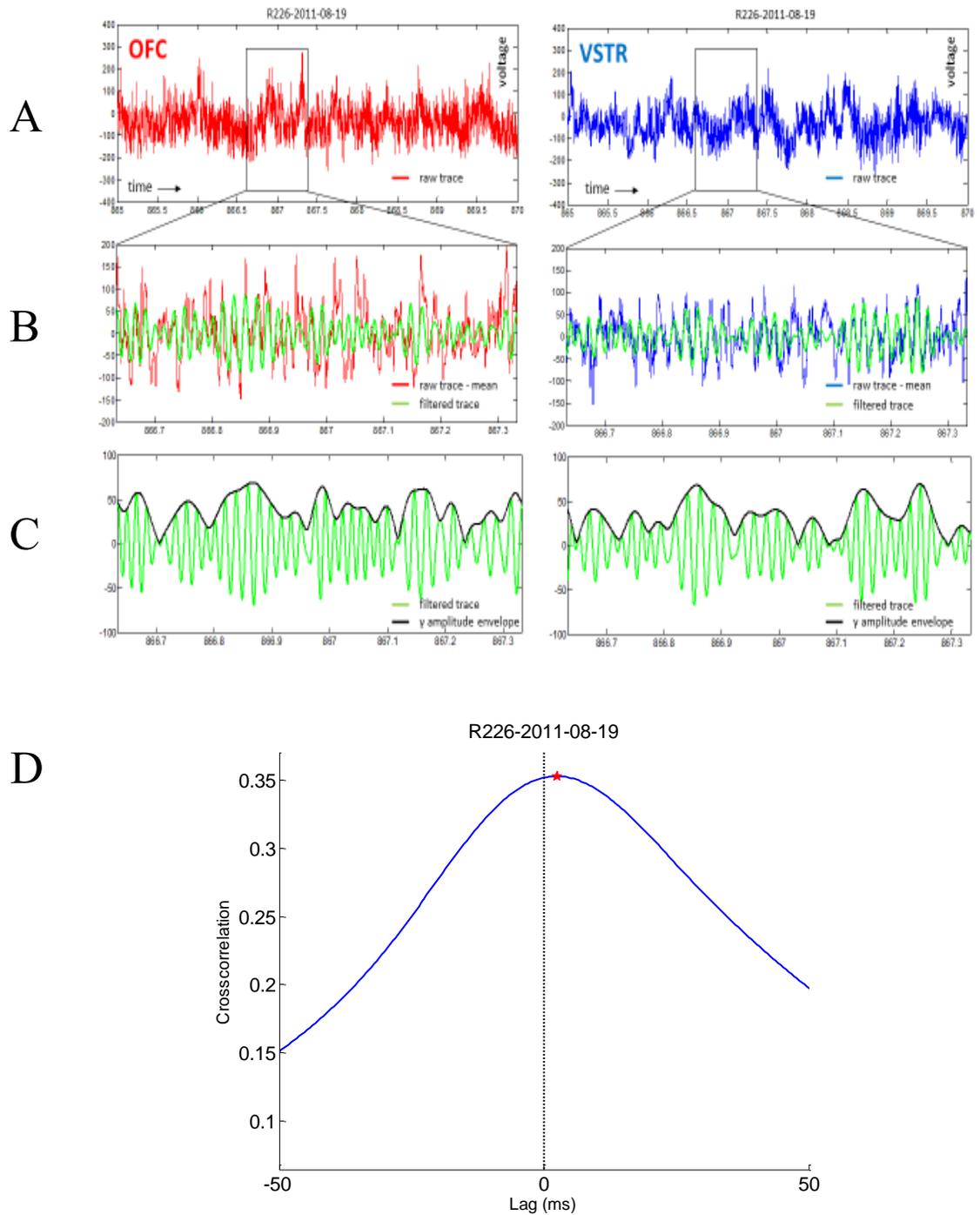


Figure 2-8. **Example of LFP data showing the cross-correlation technique.** (A) Raw voltage traces in red (OFC) and blue (vStr) during a sample LFP trace from the delay-discounting task. (B) Zoomed in to make the gamma oscillations apparent. The raw signal is filtered (here, from 40-70 Hz) and is shown in green. (C) The amplitude envelope of the filtered signal is calculated (shown in black), and these two time series are correlated. (D) Cross-correlation between times series for

the whole session, R226-2011-08-19. Red star indicates the maximum value of the correlation function, which is used to estimate of the lag between structures. Positive lags correspond to vStr leading OFC, and negative lags correspond to OFC leading vStr. Here, the max crosscorr value corresponds to a lag of 2.5ms, with vStr leading OFC.

## Chapter 3: Single-Unit and Decoding Analyses

### Introduction

The OFC and vStr are both activated by the anticipation and receipt of reward, as measured across a range of recording techniques [32, 52, 174]. Functional magnetic resonance imaging (fMRI) studies have found a high degree of overlap in the blood oxygen level dependent (BOLD) signals in these two structures, activity that scales with the expected value of reward [8]. However, the temporal resolution of fMRI is relatively slow (on the order of seconds) [175], and cannot determine the temporal ordering of OFC and vStr activity at a fine time scale. Since internal cognitive events can take place very quickly [114], it is important to understand the dynamics of reward-related activity at a fast timescale. Differences in the timing of decision-making signals in OFC and vStr—at a second or subsecond level—could reveal fundamentally distinct computations. Therefore, we made simultaneous neurophysiological recordings in the OFC and vStr of rats with thin wire electrodes, which offer high temporal resolution.

### Behavioral methods

Rats were trained to perform an economic decision-making task, the *spatial adjusting delay-discounting task* [112] (see Chapter 2: Methods), or the “DD task” for short. Briefly, rats ran on a T-maze task in which one of the choice arms (the non-delay side) offered an immediate, small reward, whereas the other choice arm (the delay side) offered a large reward after a variable delay—called the “adjusting delay.” Successive laps to the delay side increased the adjusting delay by one second, and successive laps to the non-delay side decreased the adjusting delay by one second. The initial value for the adjusting delay was chosen pseudorandomly from a distribution of one to thirty seconds. After familiarizing rats with the task, rats ran a “behavior-only” sequence for thirty days, using a custom-made backpack with an LED light to track their position. Each session lasted for

100 laps, or 60 minutes, whichever came first. In practice, rats almost always completed 100 laps within the 60 minute time interval.

Once the behavioral sequence was complete, rats underwent stereotactic surgery and were implanted with “dual-bundle” hyperdrives—using standard procedures (see Chapter 2: Methods)—that targeted the lateral OFC and the ventral striatum, with an additional auxiliary electrode placed into the hippocampus. The results in this chapter focus solely on the neural recordings made from OFC and vStr. The relationships between vStr and OFC and their relationships to the hippocampus are discussed in Chapter 4. During recovery from surgery, tetrodes were manually lowered from the brain surface toward their intended targets, with arrival at the proper depth timed to coincide with recovery from surgery and the beginning of neural recordings. As soon as rats had recovered from surgery and met the criteria for “retraining” on the task, they were plugged into the recording hardware and ran up to 30 daily recording sessions (average = 28 recording sessions), in the same manner as described above.

### Behavioral results: Discounting

During the behavior sequence, and importantly, during the recording sessions, rats showed robust “titration” behavior. During “titration” rats moved the initial adjusting delay to a stable level, at which point their behavior changed to alternation between the two maze arms, keeping the adjusting delay relatively constant for the remainder of the session. The task design aimed to elicit this type of behavior; namely, that rats would show their “indifferent point”—the point at which the value of both choice options was roughly equal—by a switch from laps to the same side to predominantly alternation laps. Figure 3-1 shows example behavior for the spatial delay discounting task.

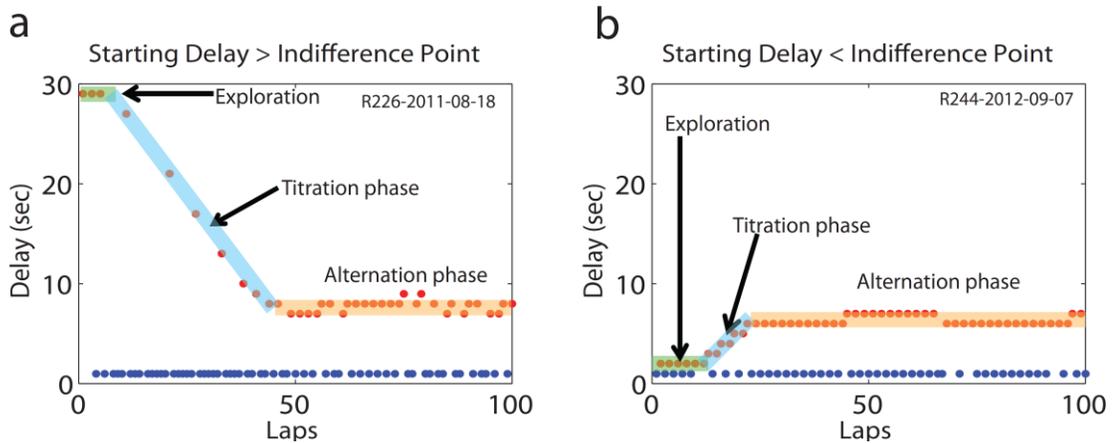


Figure 3-1. **Behavior on the delay discounting task reflects distinct phases.** Example sessions in which the initial delay (randomly chosen) is greater than (a), or less than (b), the rat’s preferred delay (i.e., his indifference point. Red dots indicate the adjusting delay on laps in which the rat chose the delayed side. Blue dots indicate the non-adjusting delay (always 1 second) on laps in which the rat chose the immediate reward side. In both sessions, initial laps were characterized by an Exploration Phase (green shading), in which the rats typically alternates sides [112] This was followed by the Titration phase (blue shading), in which rats change the adjusting delay (with successive laps to non-delay s or delay side—a & b, respectively). Rats then showed a prominent alternation pattern of behavior, termed the Alternation phase (orange shading), or the Exploitation phase—reflecting the fact that rats have found their indifference point and are “exploiting” that knowledge to continue with an optimum strategy. Figure image from [144], reprinted with permission.

Figure 3-2 displays the observed behavior on the delay-discounting task. This data demonstrates that rats showed robust titration behavior, moving the indifference point consistently to an average delay of 5-6 seconds. Thus, rats were sensitive to the two variables manipulated here—delay to reward, and reward magnitude. Rats modified their behavior according to the discounted value of the reward options.

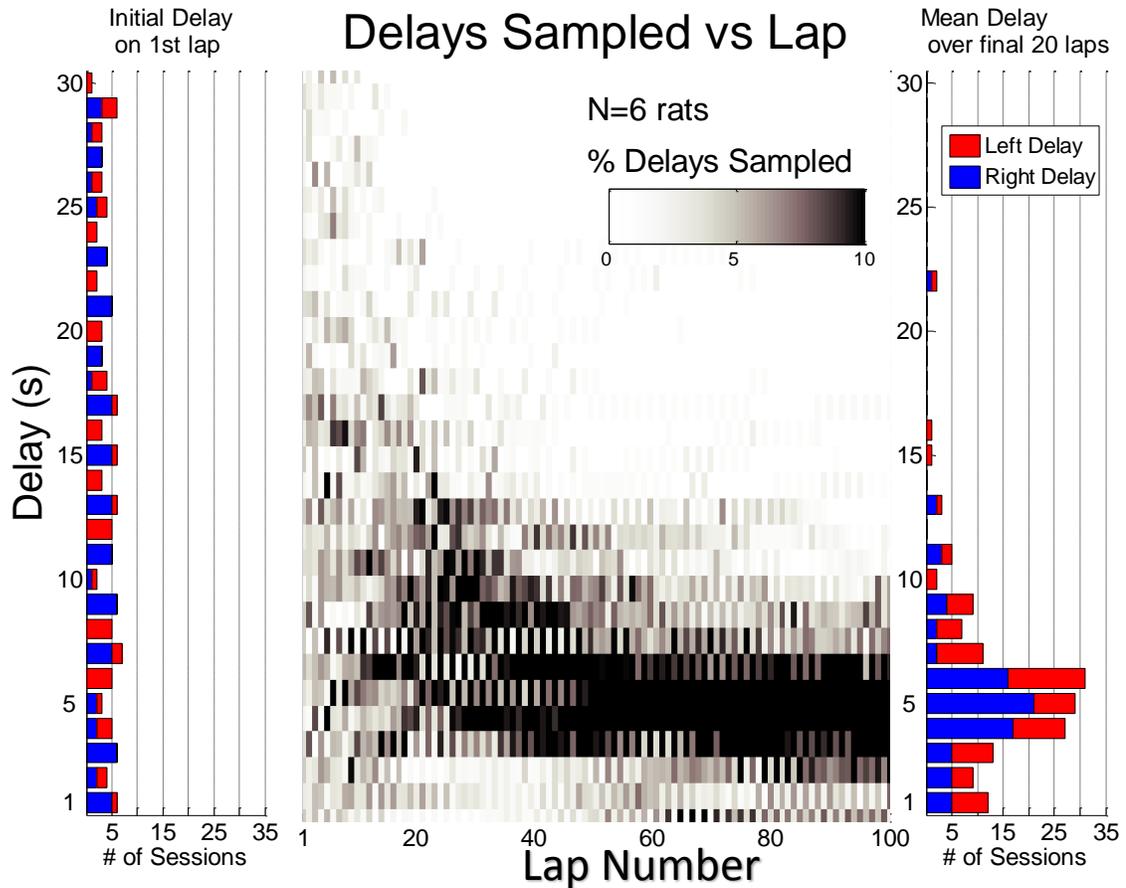


Figure 3-2. **Titration behavior on the spatial adjusting delay-discounting task.** Behavioral data from all rats ( $n = 6$  rats, 164 sessions) showing the change in adjusting delay as a function of lap within the session. In the center panel, shading represents the percentage of sessions in which the rat experienced a particular adjusting delay on a given lap. Darker shading indicates higher probabilities. Note that the sampled adjusting delays begin broadly over the entire distribution of delays from 1 to 30, but over time the animals move the adjusting delay steadily toward a preferred delay of about 5 seconds by the end of the session. The histograms on the left and right show the distribution of initial and final delays (averaged over the last 20 laps), respectively. Red indicates that the left feeder was the delayed feeder on that day, and blue indicates that the right feeder was the delayed feeder on that day. The distribution of initial delays is flat, as they were chosen by the experimenter to evenly account for all of the delays from 1 to 30. The distribution of final delays was a roughly normal distribution, centered at 5-6 seconds. Analysis is the same as in Fig. 4 from [112], but using only the rats that I recorded from.

The success of this task in eliciting delay-discounting behavior is notable for a couple of reasons, offering benefits that are unique to this type of task. First, as a spatial task, it facilitates the use of spatially defined neural decoding techniques [155] to measure cognitive processes [114, 156, 176]. Most delay-discounting tasks in rodents use operant chambers with two levers, spaced closely together. This feature of operant chambers is not a disadvantage per se, but it does make it difficult to apply spatial decoding analyses.

Second, the spatial delay-discounting task provides a consistent indifference point within subject, across days. In the traditional adjusting delay task, subjects repeat blocks of four trials each—two forced choice trials followed by two free choice trials. If the response on the two free choice trials are identical, the delay moves up or down on the next choice block (based on whether the subject chose the delayed or the immediate reward, respectively). A careful look at the behavior on adjusting delay tasks in general indicates that the adjusting delay for individual subjects oscillates substantially over time (see [177], Figure 3.5). This oscillation spans days, and therefore is not a within-session effect. Using an adjusting delay task, Torres et al. [178] found that the indifference delay (i.e. the indifference point) showed a frequency of oscillation of between 30 to 100 trial blocks (mean = 78) [178]. Slow fluctuations are prominent in other datasets as well [179] (Figure 2 of that paper).

Cardinal et al. [179] used computational simulations to model their rat subjects' actual choice behavior based on different decision rules. This analysis showed that the observed behavior could not be readily distinguished from delay-independent decision rules—including a random walk rule [179]. For the animals' observed behavior and for the simulations, the indifference delay varied substantially across sessions. On the spatial delay-discounting task, the indifference point did not vary significantly across sessions [112]. Therefore, the spatial adjusting delay-discounting task may offer a more reliable measure of animals' sensitivity to delay, in addition to being more amenable to neurophysiological recording.

Behavioral results: VTE

A prominent feature that stood out immediately in the spatial delay-discounting task is that rats displayed clear instances of vicarious trial and error (VTE)—looking back and forth at the decision point before making their choice. For a description of how VTE was quantified, see Chapter 2; Methods. The frequency of VTE events as a fraction of all laps was 13% (2099 of 16000 laps). Figure 3-3 shows the distribution of VTE and non-VTE laps across all sessions.

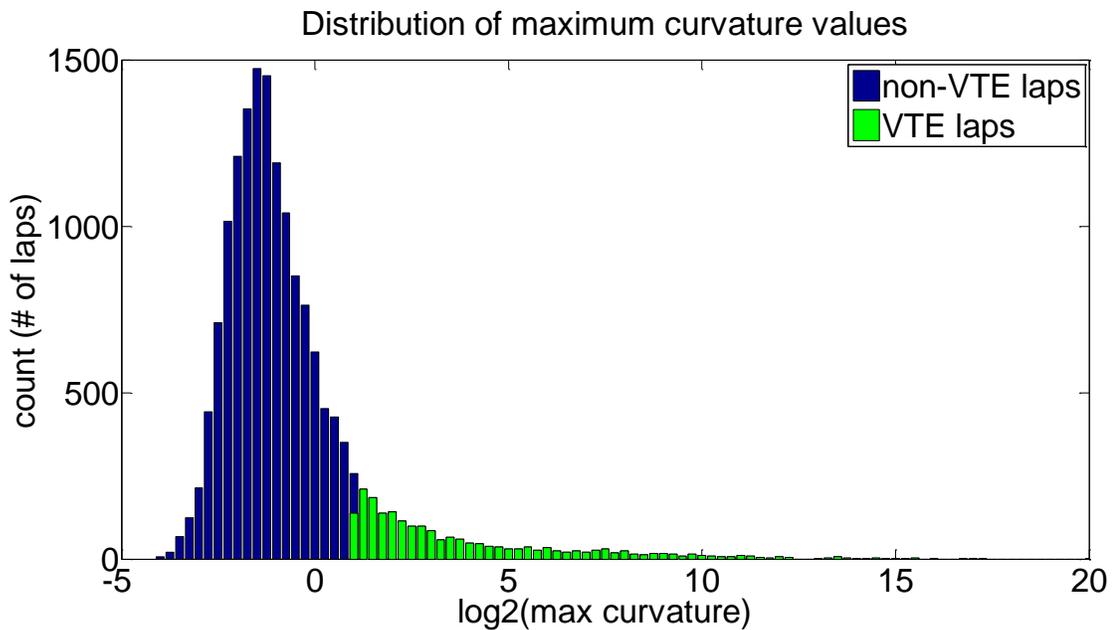


Figure 3-3. **Distribution of VTE events as quantified by path curvature.** Histogram of maximum curvature values from all laps. The degree of pause and look behavior (VTE) was quantified with a metric called curvature (see Chapter 2: Methods). High curvature values indicate pause and look behavior at the choice point. Choice point passes with a maximum curvature value greater than 2 were classified as VTE laps. Figure image from [144], reprinted with permission.

VTE occurred most often during the titration segment of the task, when animals were changing the adjusting delay, and decreased during the exploitation phase, when the adjusting delay was held constant (Figure 3-4 C & [112]). This data fits with others studies showing that VTE is

more likely when some aspect(s) of the task is changing—like after a rule reversal [83, 120, 180] or a change in outcome value [82], and that VTE behavior is an indication that the animal is considering different options in the face of changing circumstances [69].

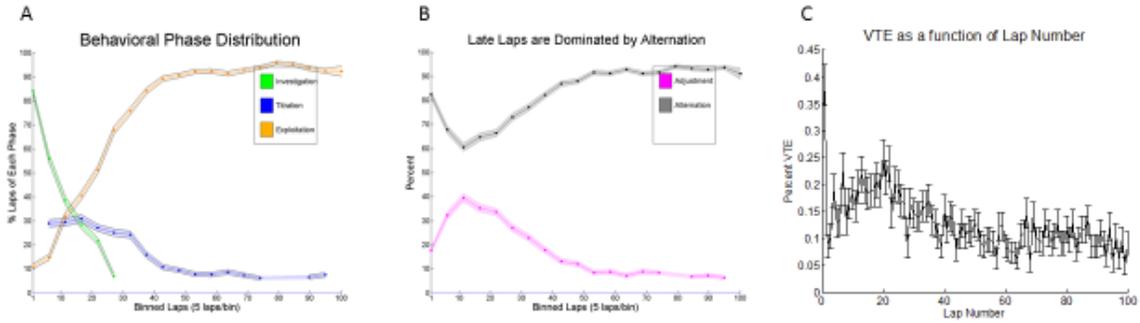


Figure 3-4. **Relationship between task phase, lap type, and VTE.** At the macro level, behavior on the DD task can be described by three distinct task phases: investigation, titration, and exploitation. Panel (A) shows the distribution of behavior phases on the DD task. The definitions for each phase are identical to that in [112]. Investigation is present in early laps and falls off steeply. The titration phase is most prevalent in early laps (10-30) and gives way to exploitation behavior (i.e. alternation) during the second half of the session. This transition can also be seen in the percentage of adjustment and alternation laps (B). Adjustment laps (returning to the same side) are relatively high early in the session, whereas late laps are dominated by alternation behavior (going to the opposite side). The frequency of VTE laps matches this pattern (C). VTE laps (defined by curvature values  $>2$ ) are most prevalent during the titration phase of the task and then fall to a baseline level during the second half of the session. Analyses in (A) and (B) are the same as in Fig. 5 a & b from [112], but using only the rats that I recorded from ( $n = 6$  rats, 164 sessions).

VTE occurred most frequently during the titration segment of the task. However, the titration segment can include both types of laps: (1) a lap in which the rat returns to the same side as on the previous lap—called a “repeating” or “same-side lap,” and (2) a lap in which the rat goes to the side not visited on the previous lap—called an “alternation lap.” Repeating laps change the adjusting delay. Alternation laps keep the adjusting delay constant. Looking at the distribution of repeating laps and alternation laps in relation to VTE, we found that VTE events were evenly

distributed between alternation and repeating laps (proportion of VTE laps that were alternation laps = 0.51), whereas non-VTE laps were primarily alternation laps (proportion of VTE laps that were alternation laps = 0.86) (Figure 3-5). This means that during non-VTE laps, simply knowing where the rat was coming from mostly predicted where the rat was going to. On the other hand, during non-VTE laps, simply knowing where the rat was coming from provided no information about where the rat was going to. Therefore, the behavioral data indicate that when the rats engaged in VTE, they did not know beforehand which side they were going to choose. This suggests that VTE occurred during moments of deliberation. When rats did not engage in VTE, they seem to have already committed to a decision—favoring alternation, those laps being predominant late in the session (when habitual behavior would be expected). Thus, the spatial adjusting delay-discounting task elicits periods of both deliberative and habitual behavior within single sessions, allowing for the comparison of neural activity during these contrasting behavioral modes.

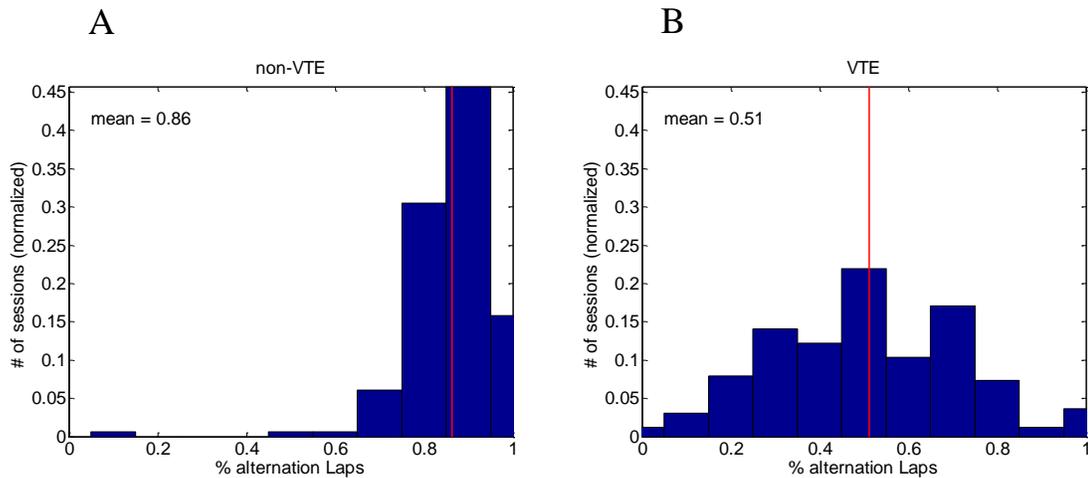


Figure 3-5. **Relationship between VTE events and lap type.** To determine the relationship between VTE events and alternation/non-alternation laps, we calculated the likelihood that a lap was an alternation or non-alternation lap given that the animal showed VTE or not on that lap. The histograms show the probability that a lap was an alternation lap given that the animal did not show VTE on that lap (non-VTE, **A**) or that the animal did show VTE (**B**). Non-VTE laps tended to be alternation laps (mean percentage alternation = 0.86), while VTE laps were as likely to be

alternation as not (mean percentage alternation= 0.51). Figure image from [144], reprinted with permission.

### Neural results: Single-unit data

One of the primary factors that influences neural activity in OFC and vStr is reward receipt. In order to characterize reward responsivity of neurons in OFC and vStr, we calculated the firing rates of neurons in these two structures in a window from 0 to 4 seconds after feeder trigger events. See Chapter 2: Methods, for a description of how reward responsive units were defined. The terms “feeder trigger” and “feeder fire” are used synonymously here. This time window encompassed a short period of reward anticipation (as the pellet(s) descended from the pellet dispenser into the feeder port) and the approximate time of reward consumption. An examination of peri-event time histograms (PETHs) aligned to feeder fire events (both feeders taken together) showed that neurons in both OFC and vStr showed robust firing rate changes in response to reward receipt. This firing rate modulation included both increases and decreases in firing rate as compared to the session average, for both OFC and vStr (see Figure 3-6, panels a-d for examples). Figure 3-6 (panels e-f) shows the distributions of reward-responsive units in vStr and OFC, according to their z-scored firing rate during the reward window. A large fraction of cells modulated their firing rate around the time of reward delivery: 70.2% (665/947) of neurons in vStr and 68.5% (1201/1754) of neurons in OFC showed a significant change in firing rate, with the fraction of reward-modulated cells split evenly between those that increased firing rate and those that decreased firing rate. The percentages of reward-responsive units reported here are consistent with other reports in the literature [48, 54, 120, 181, 182], and they are in agreement with the primacy of reward processing that is attributed to the OFC and vStr.

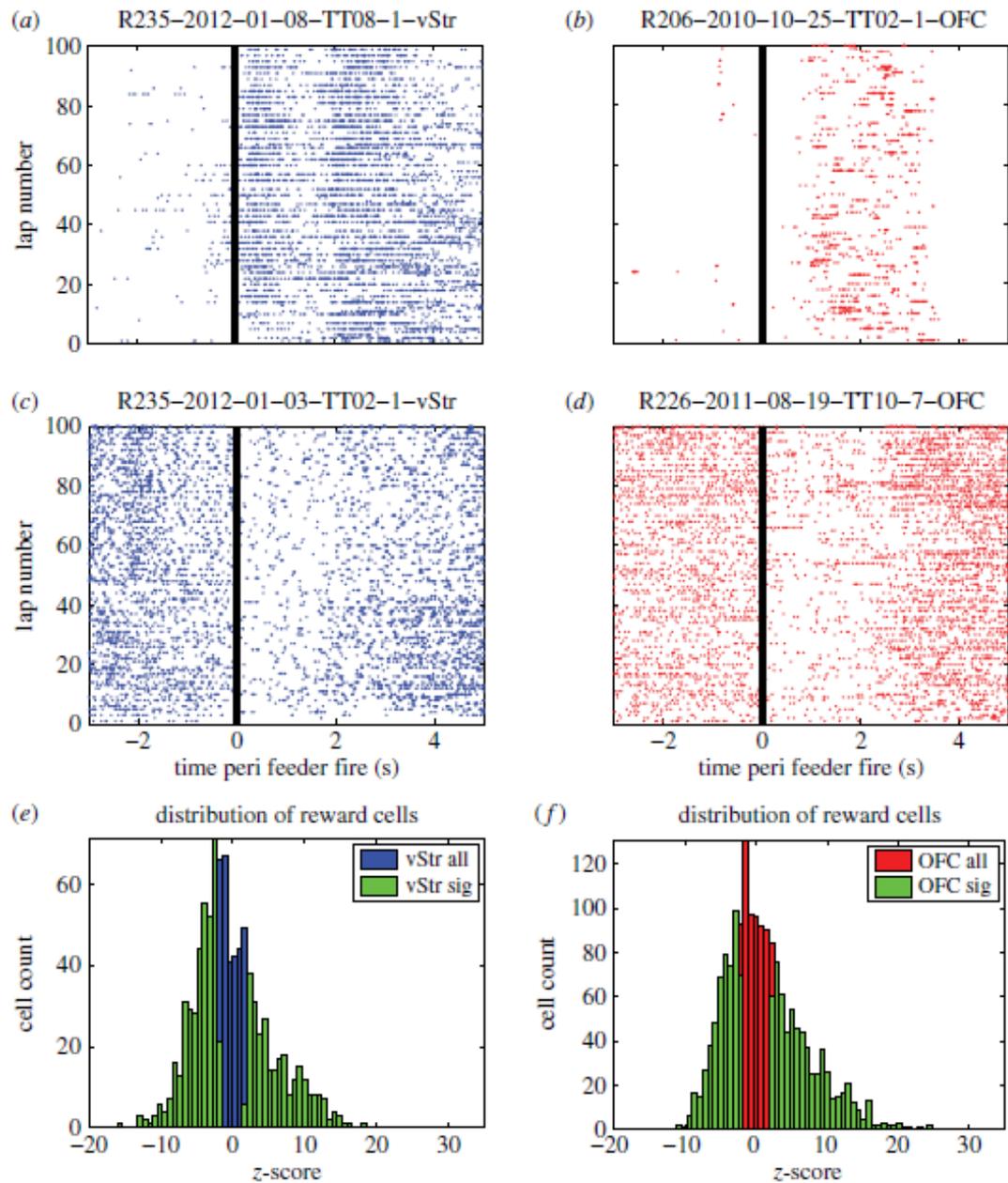


Figure 3-6. **Reward responsive units in vStr and OFC.** Example single unit recordings from vStr (a) and OFC (b) that showed a significant increase in firing rate around the time of reward receipt. Other cells showed a significant decrease in their firing rate around the time of reward receipt (c, d). The distribution of firing rates around the time of reward receipt is shown for all cells in vStr (n=947) (e) and OFC (n=1754) (f) in blue and red, respectively. Cells that significantly changed

their firing rate (z-test,  $p < 0.05$ ) compared to a bootstrap distribution (see Chapter 2: Methods) are overlaid in green. Figure image from [144], reprinted with permission.

Theoretical accounts of OFC and vStr function also emphasize their role in coding reward *value*. These accounts make several testable predictions regarding the firing rates of cells on the delay-discounting task. If single cells are coding the value of reward, then they should (1) change their firing rate in the same direction to both rewards (consistency), (2) change their firing rate *more* for the larger reward (magnitude coding), and (3) modulate their firing rate according to the delay to reward (temporal discounting). Each of these predictions is taken up in turn below.

A large fraction of neurons significantly changed their firing rate during reward receipt, with different populations of cells changing their firing rate to reward in different directions (Figure 3-4 e,f). Although many reports only analyze cells that show firing rate increases to task-related events, we consider decreases in firing rate to be equally valid instances of information coding. Therefore, we included both populations of cells in our analyses. The high percentage of reward-responsive units indicates that reward receipt is a strong factor in modulating firing rates on this task. Interestingly, cells in OFC and vStr did not use a unitary coding scheme (i.e. some cells increasing to reward, some cells decreasing). However, this latter assertion requires an examination of each reward response separately, because the magnitude of reward differed at the two feeder sites.

If neurons in OFC and vStr are indeed coding reward, then they should show firing rate changes *in the same direction* to rewards at both feeder sites. We term this *consistency*. In order to visualize differential firing to the two rewards, we plotted the average (z-scored) firing rate to the small reward versus the average firing rate to the large reward for each neuron (Figure 3-7, panels a,b). Reward-responsive cells are shown in color (non-responsive cells in black). In these plots, consistency of reward coding is indicated by colored points that fall within the upper right and bottom left hand quadrants (increasing or decreasing to both rewards, respectively), versus points

that fall within the upper left and bottom right hand quadrants (increasing to the large reward but decreasing to the small reward, or vice versa). Among reward-responsive cells, a significantly larger fraction of these neurons coded for reward in the same direction as opposed to coding reward in opposite directions (binomial test,  $p < 0.05$ ) (Figure 3-7 e,f, left side bar graphs). Thus, the majority of neurons in vStr and OFC exhibited consistency in their reward coding.

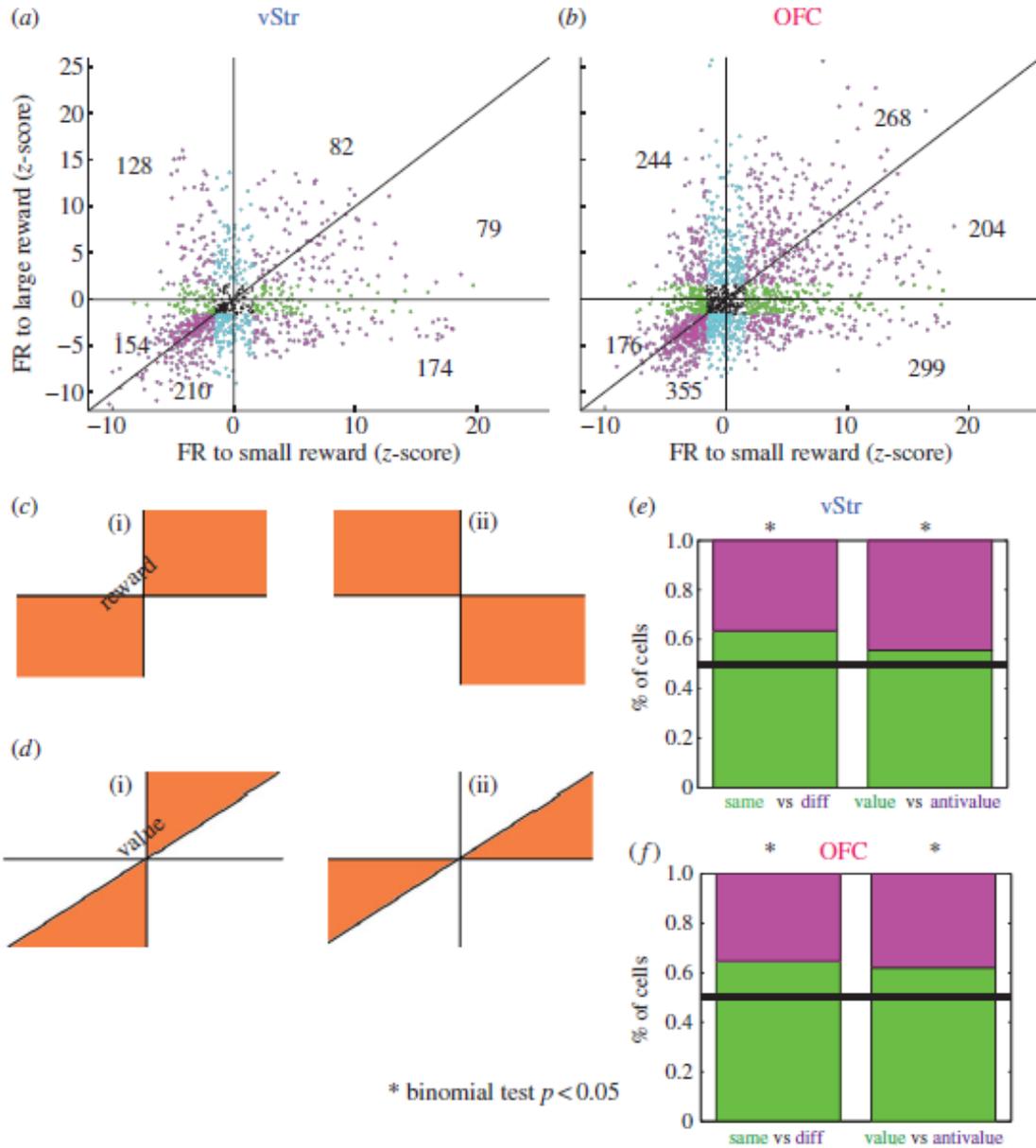


Figure 3-7. **Single cells code for reward and for value.** Firing rate in response to reward at each feeder location for all vStr cells (n=947) **(a)** and all OFC cells (n=1754) **(b)**. Cells that showed a significant change in firing rate are displayed in color. Significance was determined by comparing reward responses to a bootstrapped distribution (see Chapter 2: Methods). For significance, blue=small reward, green=large reward, magenta=both rewards (cell counts for all blue, green, or magenta cells are shown within each of the six sectors), and black=non-significant. **(c)** Diagram showing what reward coding should look like. Axes as in (a,b). If single cells are responsive to reward, they should either increase their firing rate at both feeders, or decrease their firing rate at both feeders (i.e. a *consistent* response) **(c(i))**: “same”, upper right and lower left quadrants). Firing rate changes in the opposite direction are incompatible with reward coding **(c(ii))**: “diff”, upper left and lower right quadrants). **(d)** Diagram showing what value coding should look like. Axes as in (a,b). If neurons code not only for reward in general, but also the magnitude of reward, then changes in firing rate should be greater for the three pellet reward than for the one pellet reward **(d(i))**: “value”,  $|large| > |small|$ ). Firing rate changes in the opposite direction are incompatible with value coding **(d(ii))**: “anti”,  $|small| > |large|$ ). **(e,f)** The left side of each bar graph shows the percentage of reward responsive cells classified as reward coding (“same” vs. “diff”). The right half of each bar graph shows the percentage of cells that demonstrated value coding (“value” vs. “anti”). In both structures, significantly more cells coded for reward than not and for value than not (binomial test,  $p < 0.05$ ). Figure image from [144], reprinted with permission.

Numerous reports have shown that both the BOLD signal in OFC and vStr and ensembles of cells in these structures code for reward magnitude [8, 16, 37, 49]. In order to determine whether single units coded for the magnitude of reward on the delay-discounting task, we used the same plot as above (Figure 3-7). If cells code for reward magnitude, then they should show a greater change in firing rate to the larger magnitude reward. This would manifest itself as points that fall between the ordinate and the identity line, as opposed to points that fall within between the abscissa and identity line (see Figure 3-7, panel d). Among cells that showed *consistency* in their reward coding (Figure 3-7 c(i)), significantly more of these cells showed magnitude coding than those that did not (binomial test,  $p < 0.05$ )(Figure 3-7 d,f, right hand bar graphs).

While magnitude coding was more prevalent than not (61% and 55% for OFC and vStr, respectively), these effects were not overwhelming. The absolute number of cells that had significant changes in firing rate for *both* feeders and fell within the “value” sectors (Figure 3-6 d(i)) was 410/1754 (23.4%) for OFC and 181/947 (19.1%) for vStr. These percentages are in line with other studies [37, 49]. Altogether, the results on consistency and magnitude coding are consistent with the literature, which shows that reward coding is significant, but does not explain most of the variance in activity in these structures. Firing rates in the OFC and vStr can be influenced by a number of different task variables in addition to reward magnitude, including sensory, motor, and contextual variables [43, 45, 46].

Accounts of OFC and vStr function that emphasize the *integration* of different value parameters would predict that neural firing should be influenced not just by reward magnitude, but by the delay to reward as well. To investigate this possibility, a linear regression analysis was conducted for firing rate as a function of delay to reward. Because the delay to reward at the one-pellet feeder was fixed at one second, this analysis was restricted to laps to the adjusting delay side. If neurons engage in temporal discounting of reward value at the time of reward receipt (using the same 0-4 sec window), then they should show a negative relationship between the firing rate and the “just-experienced” delay. In other words, firing rate should decrease with increasing delays. In this case, the beta weight for the linear regression would be negative. One such example cell is shown in Figure 3-8. Alternatively, cells could show a significant regression slope, but in the opposite direction. This would indicate delay coding, but in a direction opposite to that typically expected for temporal discounting. A significant slope in either direction we term “retrospective coding,” because the firing rate modulation is correlated with what has already happened.

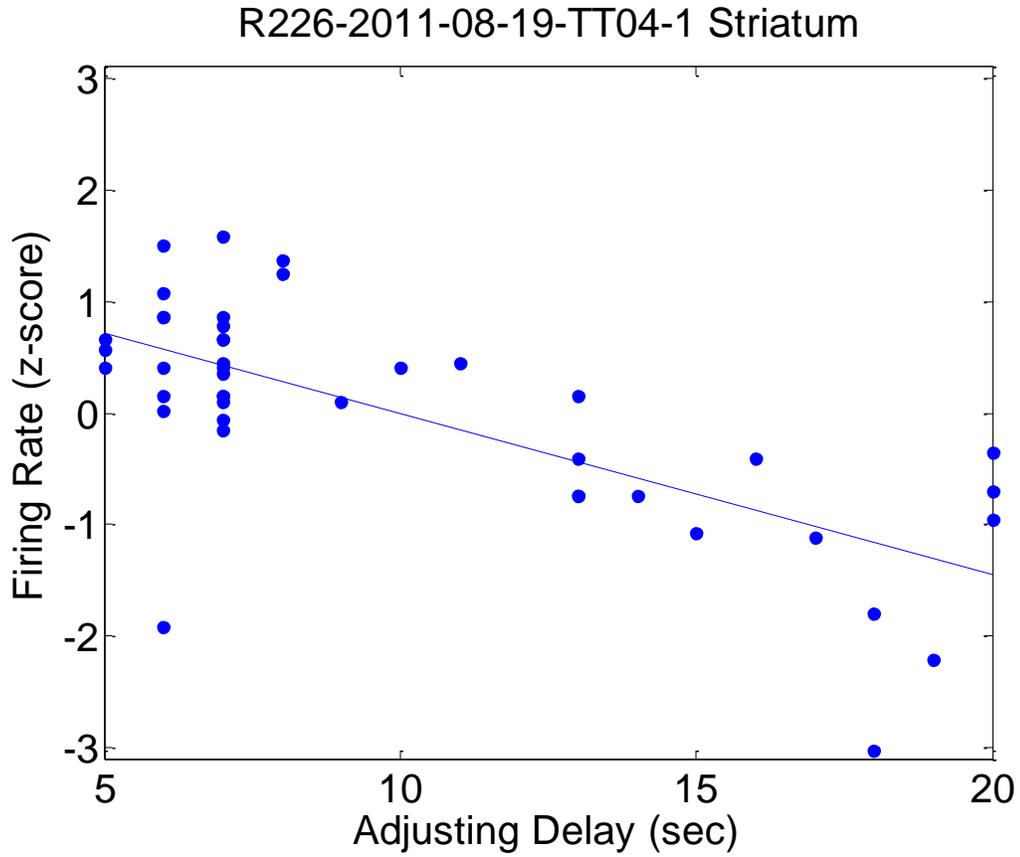


Figure 3-8. **Example “delay-discounting” neuron.** During the reward receipt window, this vStr cell decreased its firing rate with increasing delays.  $R^2 = 0.5$ . Slope ( $\beta$  weight) =  $-0.15$ ,  $p < 10^{-5}$ .

Most studies of temporal discounting examine discounted value at the time that cues are presented indicating the choices, or at the time of action selection. In order to examine temporal discounting during reward anticipation, we also regressed the firing rate for the first one second of the adjusting delay against the adjusting delay for that lap. Again, a negative slope here would be consistent with temporal discounting—reduced responding for longer delays. A significant positive slope would still indicate delay coding, but in the opposite direction. Both of these we term "prospective coding", because the firing rate modulation is correlated with the “to-be-experienced” adjusting delay.

The distribution of beta weights for delay coding in vStr and OFC is shown in Figure 3-9. During the first second of the adjusting delay (prospective coding), 288/1554 (18.5%) and 196/846

(23.2%) of cells in OFC and vStr, respectively, had significant regression slopes<sup>10</sup>. During the reward receipt window (retrospective coding), 376/1733 (21.7%) and 271/907 (29.9%) of cells in OFC and vStr, respectively, had significant regression slopes. Both of these distributions were significantly different between brain regions ( $X^2$  test,  $p < .01$ ). We ran a shuffled control and found that across many iterations, about 5% of cells came out as significant when we randomized the firing rates to the delays (i.e. shuffling the firing rates with respect to lap number within each neuron). Thus, between 20-30% of cells showed some form of delay coding. This coding was more pronounced in the vStr than in the OFC, and it was stronger for retrospective than for prospective coding, although these differences were relatively modest.

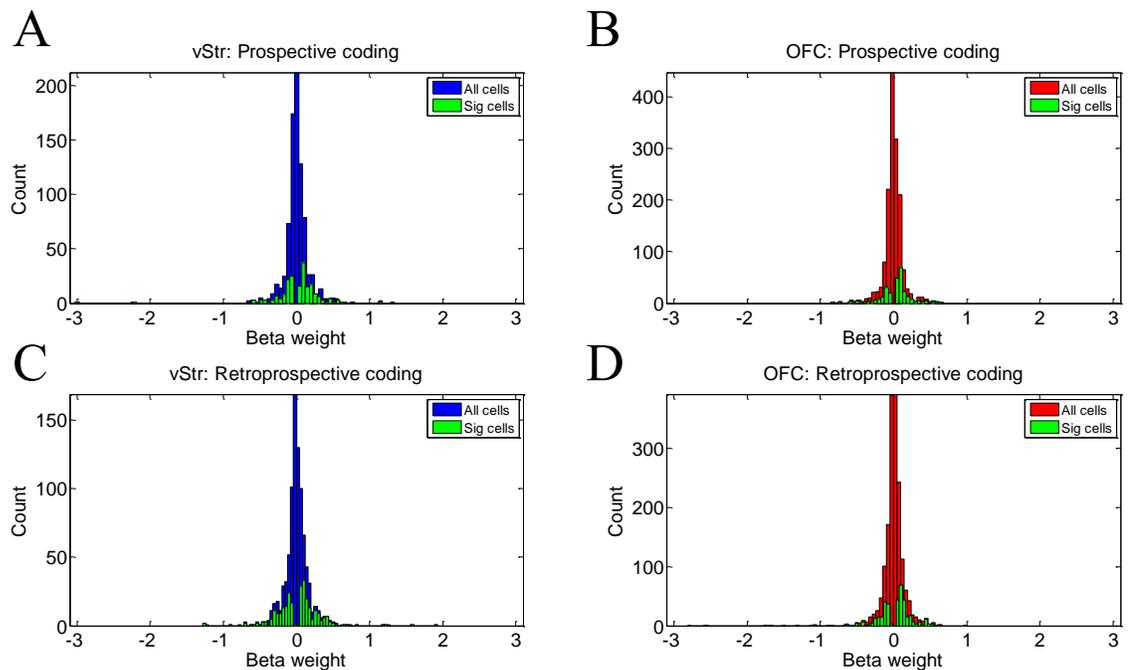


Figure 3-9. **Distribution of delay-coding cells on the DD task.** Each plot shows the distribution of beta weights for all neurons in in vStr (blue) and OFC (red) with significant cells (slope not equal to zero, linear regression,  $p < .05$ ) displayed in green. Beta weights  $> 0$  indicate a firing rate increase with increasing adjusting delays. Beta weights  $< 0$  indicate a firing rate decrease with increasing

<sup>10</sup> Some cells were excluded because they did not have any spikes during the event windows for at least two different adjusting delays. Therefore, the denominator may differ slightly from the total number of recorded cells (OFC = 1754, vStr = 947).

adjusting delays. For prospective coding, firing rates were taken from the first 1 second of the adjusting delay. For retrospective coding, firing rates were taken from the [0 4] sec window after feeder fire.

vStr Prospective (+ $\beta$ ) 114/196 (58.2%)	vStr Prospective (- $\beta$ ) 82/196 (41.8%)	OFC Prospective (+ $\beta$ ) 186/288 (64.6%)	OFC Prospective (- $\beta$ ) 102/288 (35.4%)
vStr Retrospective (+ $\beta$ ) 151/271 (55.7%)	vStr Retrospective (- $\beta$ ) 120/271 (44.3%)	OFC Retrospective (+ $\beta$ ) 225/376 (59.8%)	OFC Retrospective (- $\beta$ ) 151/376 (40.2%)

Table 3-1. **Cell counts for delay-responsive cells.** vStr is in blue, OFC in red. Among cells that showed significant delay coding prospectively (first row), and retrospectively (second row), the fraction of cells that had either positive (+) or negative (-) beta weights ( $\beta$ ) are shown. For both prospective and retrospective coding, the proportion of positive beta weights was greater than the proportion of negative beta weights (comparing within brain area), and these distributions were also significantly different,  $X^2$  test,  $p < .01$ .

Neuroeconomic accounts of decision-making with delays support the notion of temporally discounted value signals at the neural level [14]. Temporally-discounted value coding would predict a preponderance of negative beta weights: cells should fire less with increasing delay. This corresponds to the lower value associated with cues that predict long delays and, as some have suggested, the lower value associated with the reward itself if it was preceded by a long delay [37]. The counts for cells with significant positive and negative beta weights are shown in Table 3-1.

Our data do not support temporal discounting at the single cell level on the DD task. For each condition (prospective and retrospective), and across brain regions (vStr and OFC), significantly more cells had positive beta weights than negative beta weights ( $X^2$  test,  $p < .01$ ). In other words, cells were more likely to increase their firing rate than to decrease it, with increasing delay. This finding runs counter to other studies that have reported greater activity in OFC during the cue epoch [183] and during reward receipt [37] for shorter delays.

However, the interpretation of those studies is not so clear cut. In the primate study by Roesch & Olson [183], there were many more “short-preferring” than “long-preferring” cells during the cue period (15 neurons vs. 1 neuron, and 12 neurons vs. 0 neurons, for the two monkeys, respectively). However, this effect was reversed during the “post-delay period” (saccade to reward delivery) [183] (Table 1 of that paper), the reason for this being unclear. In direct contrast to the latter result, Roesch, Taylor, & Schoenbaum [37] found that more neurons in the rodent OFC were “short-preferring” during the reward receipt period (i.e post-delay) (Figure 4 of that paper). There was, however, a small but significant proportion of OFC cells that preferred the delayed reward, and they increased their firing rate progressively during the delay (Figure 6 of that paper). Thus, there is evidence for temporal discounting in OFC, but the two studies differed as to when it occurred. Our results did not show temporal discounting (in the traditional sense) for either epoch.

Of interest to theories of OFC function is whether delay preference tracks magnitude preference at the cellular level. Value coding in a common-currency would suggest that activity in single cells should systematically co-vary with delay and magnitude. On this point, the two studies mentioned above are basically consistent. Roesch & Olson [183] found a significant, positive correlation between delay and magnitude coding *during the cue period*, but not during the post-delay period (Figure 10 of that paper). Roesch, Taylor, & Schoenbaum [37] found no significant correlation during the post-delay period. Due to their experimental design, the authors of the latter paper [37] were unable to test for a correlation during the cue period. So it is at least possible that magnitude and delay coding are correlated during reward anticipation in the rodent as well.

We evaluated the correspondence between delay coding and magnitude coding by constructing correlation plots, as used in the aforementioned studies. The reward index (“larger preferring” versus “small preferring”) was calculated in the same manner, except for the fact that we used z-scored firing rates instead of raw firing rates. The formula is  $(\text{Big} - \text{Small}) / (\text{Big} + \text{Small})$ , where “Big” and “Small” are firing rates to the large and small reward feeders, respectively. Because delay and reward size were not independently varied in the DD task, we used the beta

weights for delay coding to measure delay preference. A negative beta weight indicates a “short-preferring” neuron and a positive beta weight indicates a “long-preferring” neuron.

Figure 3-9 shows the correlation between delay coding and magnitude coding for prospective (cue event) and retrospective (reward receipt) epochs. None of the conditions showed a significant correlation. In other words, there was no systematic relation between delay coding and magnitude coding at the single cell level. This data argues against an integrated value signal on the DD task, in line with the rodent data in [37]. The data in the rodent and the monkey studies by Roesch also found spatial selectivity in OFC cells, arguing against value coding in a purely abstract manner [184].

In summary, the single-unit analyses described above showed that, during reward receipt, a large fraction of OFC and vStr neurons were reward-responsive. Moreover, OFC and vStr cells showed consistency in their firing rate modulation, they coded for reward magnitude, and they showed significant delay coding. However, the data did not support the integration of delay and magnitude into a common coding scheme.

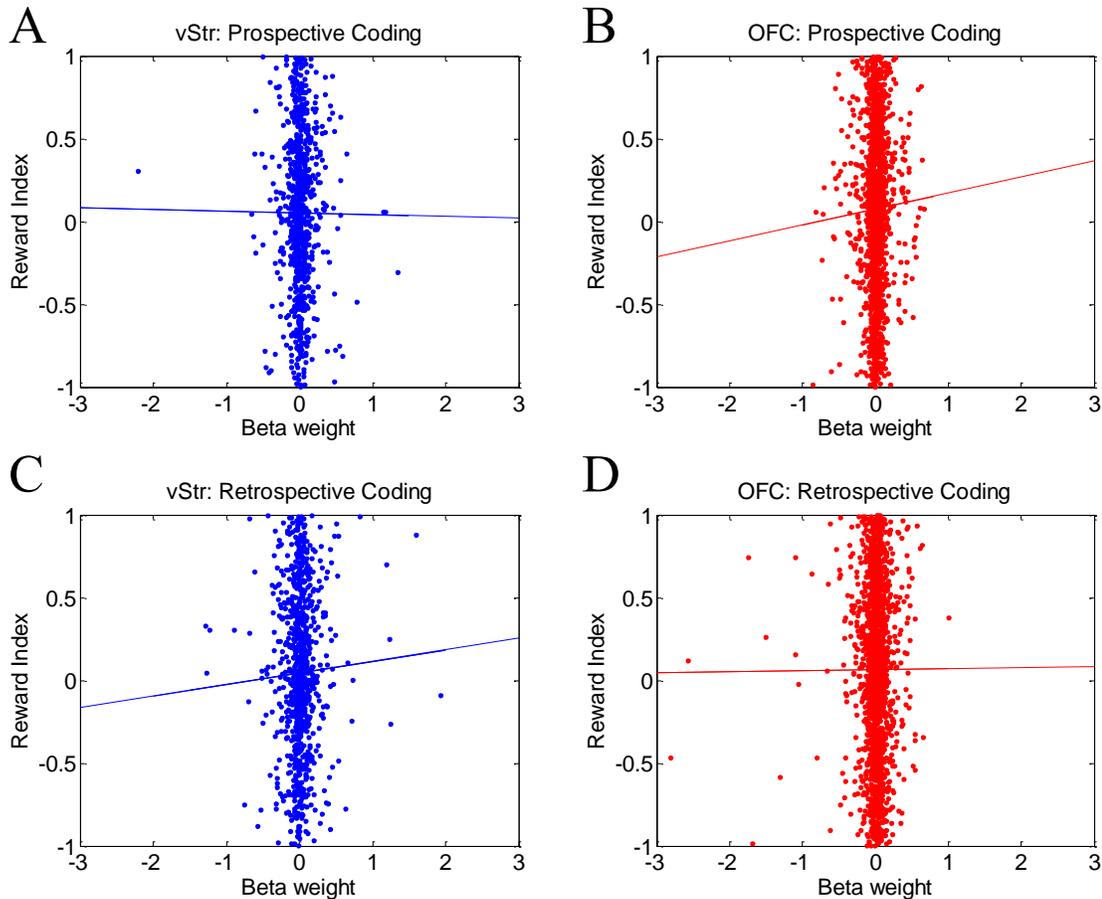


Figure 3-10. **Correlation between delay coding and reward preference.** Each plot shows the correlation between the beta weights for Prospective (A & B) or Retrospective (C & D) delay coding and the Reward Index amongst all recorded cells (vStr in blue, OFC in red). Beta weights were calculated as described above. The Reward Index was calculated using the z-scores for each cell (depicted in Figure 3-7 a,b):  $(\text{abs}|\text{Large}| - \text{abs}|\text{Small}|) / (\text{abs}|\text{Large}| + \text{abs}|\text{Small}|)$ . Values greater than 0 indicate a “large reward preferring” cell. Values less than 0 indicate a “small reward preferring” cell. Integration of delay and magnitude in a common coding scheme would manifest itself as a negative slope in these plots. Negative beta weights (“short preferring neurons”) would be associated with a positive reward index (“large preferring neurons”)—clustering in the upper left hand quadrant. Positive beta weights (“long preferring neurons”) would be associated with a negative reward index (“small preferring neurons”)—clustering in the lower right hand quadrant. None of the plots here were significant for the regression of beta weight against Reward Index ( $p$  values  $> 0.3$ ,  $R^2$  values  $< 0.005$ ).

### Neural results: Ensemble decoding

To get a fuller picture of reward coding in OFC and vStr, we used Bayesian decoding to quantify the population dynamics of reward-site representations on the DD task. Bayesian decoding offers a principled way of combining information contained in the population of cells within a given ensemble [155, 185]. As noted above, many reward-responsive cells on the delay-discounting task decreased their firing rate in response to reward. A distinct advantage of Bayesian decoding is that decreases in spiking activity also contribute information to the measure being calculated. Additionally, Bayesian decoding techniques can help reveal “covert” cognitive events that are not evident in raw firing rate measures [48, 114, 156].

Previous studies in OFC [120] and vStr [48] have shown enhanced reward-site decoding during VTE events. The fact that these value-related structures show reward representations during VTE suggests that they are both involved in the deliberation process. Comparing the data from Figure 11 in the paper by Steiner & Redish [120] and Figure 10 in the paper by van der Meer & Redish [48], the timecourse of decoding suggests that vStr precedes OFC in the representation of reward during VTE. However, these studies used different criteria for classifying VTE events and for selecting the moment of “reversal”/“reorientation”, and they were performed in separate groups of animals. Therefore, the precise timing of reward decoding during VTE remains unknown, leaving it unclear whether vStr and OFC are playing similar or different roles during the deliberative process. Simultaneous recordings from our dataset allow for a direct comparison of the timing of reward-related activity in OFC and vStr.

The methods used to measure reward-related decoding on the DD task were similar to those described previously [48, 120]. See Chapter 2; Methods, for a full description. We used a spatial decoding algorithm to measure neural representations focused on particular zones of interest—the reward sites. On the delay-discounting task, reward receipt only occurred at circumscribed

locations—a small area around the feeder port. Because reward receipt was restricted to those particular spaces on the maze, a spatial decoding algorithm—when we examine decoding to those areas—will necessarily provide information about the animal’s neural representation of reward. It is important to keep in mind that the use of a spatial decoding algorithm does not imply or depend upon OFC or vStr neurons having any spatial firing correlates. Spiking activity in OFC and vStr cells that is similar to that which occurs during reward receipt will drive the locus of spatial decoding to the reward sites [48, 120].

To isolate reward site decoding specifically, we first determined what regions of the maze the animal sampled during the 4 second period after feeder fire (same window as used above in the single unit analyses). We defined the feeder sites as those spatial bins where there was nonzero occupancy—meaning the animal was present there for at least some amount of time within the session—during the reward response window. We defined *pFeeders* as the posterior probability values for the reward sites (both sites averaged together). *pFeeders* values were normalized for the number of spatial bins (see Chapter 2: Methods, for a description). Values greater than 1 indicate higher than chance levels of decoding to the feeders. Values less than 1 indicate lower than chance levels of decoding to the feeders. The strength of feeder site representations was calculated at each moment during the session, in 250ms bins.

In order to determine whether the *pFeeders* measure tracked reward anticipation and reward receipt, PETHs were constructed aligning *pFeeders* to Zone Entry (ZE) and Feeder Fire (FF) times on the maze. Both OFC and vStr showed a dramatic rise in the feeder site representation after the cue signaling the countdown to reward delivery (Figure 3-11, a,b,c,e), with a second peak at the time of reward delivery itself (Figure 3-11, a,b,d,f). *pFeeders* remained high during the adjusting delay, demonstrating a sustained representation of the impending reward (Figure 3-11, middle and bottom panels). The dynamics of the *pFeeders* representations qualitatively matched the timecourse and shape of the population firing rates in OFC and vStr. See Figure 3-12 for

comparison. This data verifies that *pFeeders* is measuring a meaningful neural signature of the animals' reward response.

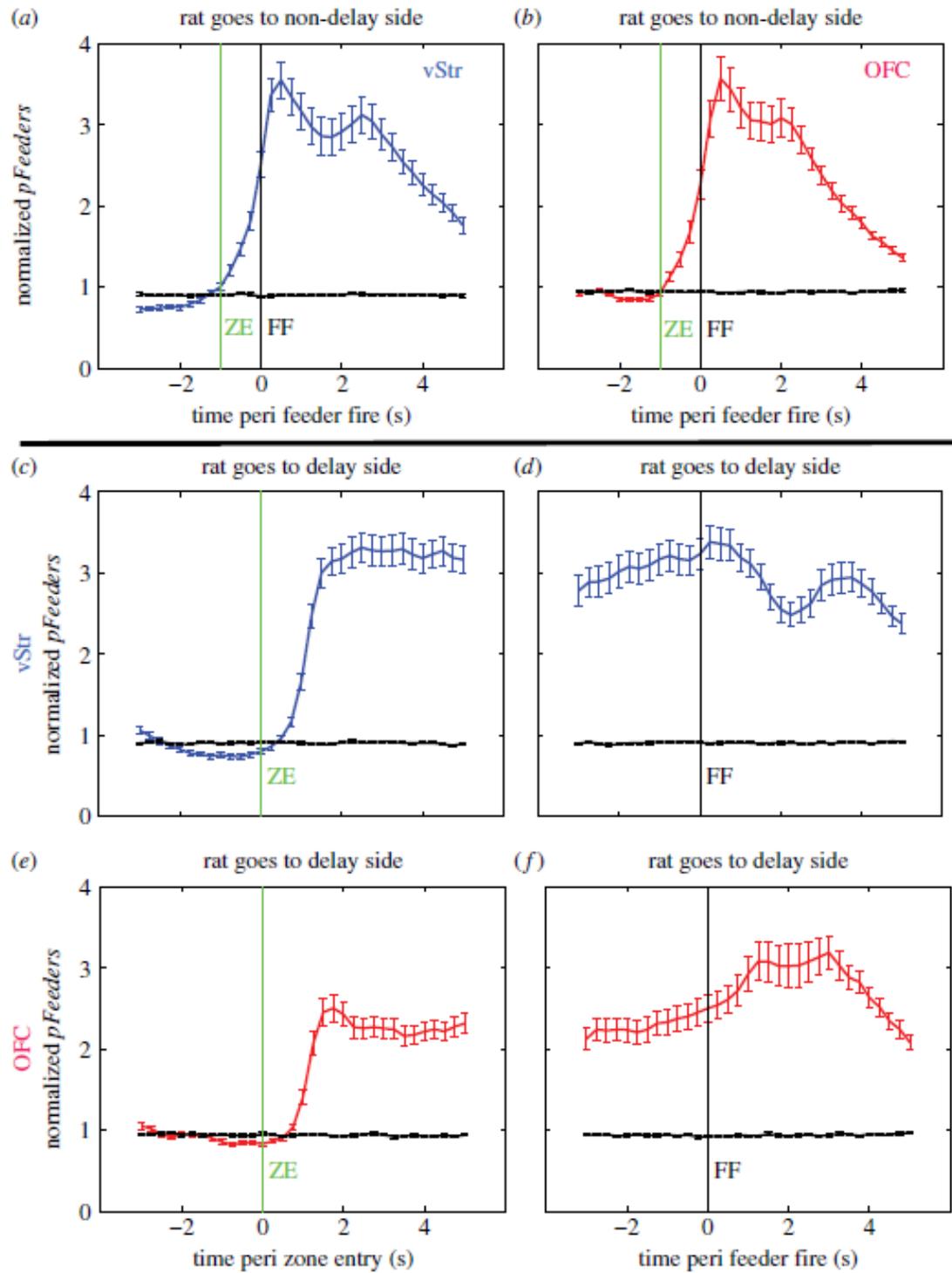


Figure 3-11. **Ensemble decoding to the reward sites.** Each row shows peri-event time histograms (PETHs) of Bayesian ensemble decoding to the feeder sites (normalized  $pFeeders$ ). Only sessions with at least five cells in each structure were included in this analysis ( $n = 85$  sessions). Lines represent the average across sessions, with error bars representing the standard error of the mean (SEM). **(a,b)** Decoding for laps when the rat chose the non-delay side. “ZE” indicates feeder zone entry (which coincides with choice point exit) and “FF” indicates when the feeder fires the pellet reward. **(c,e)** Feeder site decoding for laps when the rat chose the delay side, aligned to feeder zone entry. Decoding remained high during the delay. **(d,f)** Decoding aligned to feeder fire (FF) while the rat waited at the delayed feeder site. Note that  $pFeeders$  values were high at the end of the delay, but still peaked slightly after the feeder click and pellet delivery. Figure image from [144], reprinted with permission.

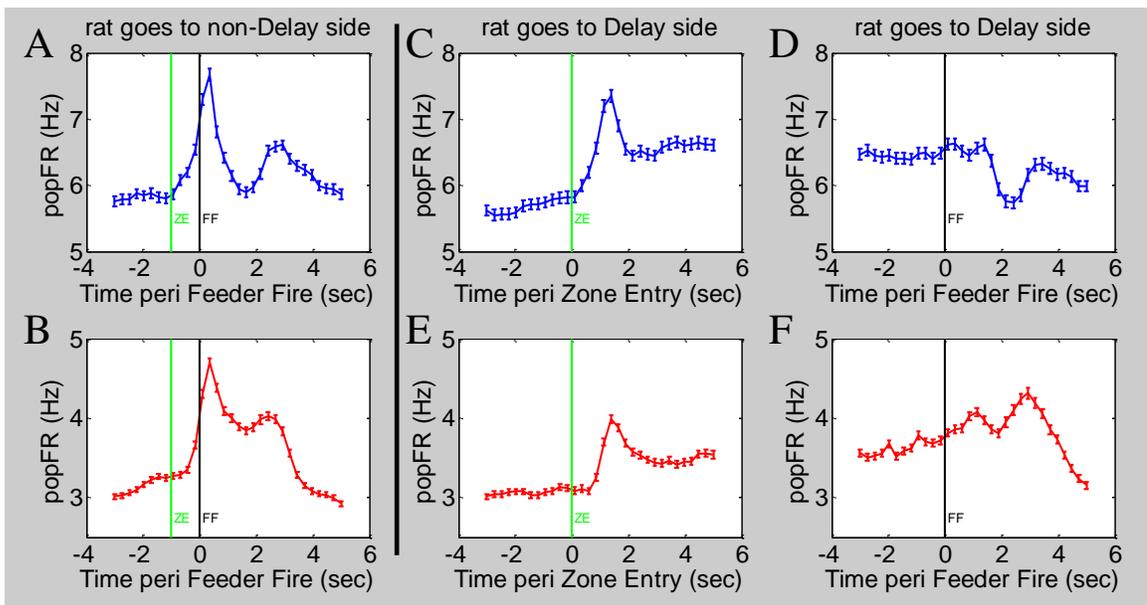


Figure 3-12. **Population firing rate PETHs.** Average population firing rate  $\pm$  SEM for vStr (blue) and OFC (red) aligned to Zone Entry (ZE) and Feeder Fire (FF) events. Includes all sessions with at least five cells in the ensemble. Conventions are as in Figure 3-11. Non-delay laps in **(A)** and **(B)**. Delay laps for vStr in **(C)** and **(D)**, and delay laps for OFC in **(E)** and **(F)**. Note the sustained activity between the second and third columns, spanning the adjusting delay.

In order to investigate the timing of reward-related representations during deliberative behavior, we compared the timecourse of  $pFeeders$  values in OFC and vStr during passes through the choice point. For VTE laps, data were aligned to the “TurnAround” point; essentially, the point

of maximum curvature in the rat’s trajectory (see Chapter 2: Methods, for how curvature was calculated). For non-VTE laps, data were aligned to the “MidPoint,” defined as the halfway point of the trajectory through the choice point zone. As shown in Figure 3-13 A, there was a substantial increase in feeder site decoding in vStr for VTE passes as compared to non-VTE passes. Importantly, this increase in decoding occurred before the auditory cues that signaled the delay (which occur upon exit from the choice point zone), and at a physical location separate from the site of reward. Thus, this increased decoding to the reward site represents a non-local or “covert” representation of future reward. Additionally, this increased decoding began prior to the time of TurnAround, before the rat turned towards the chosen side. This indicates that the increased neural representation of reward in vStr occurred before the animal made its decision.

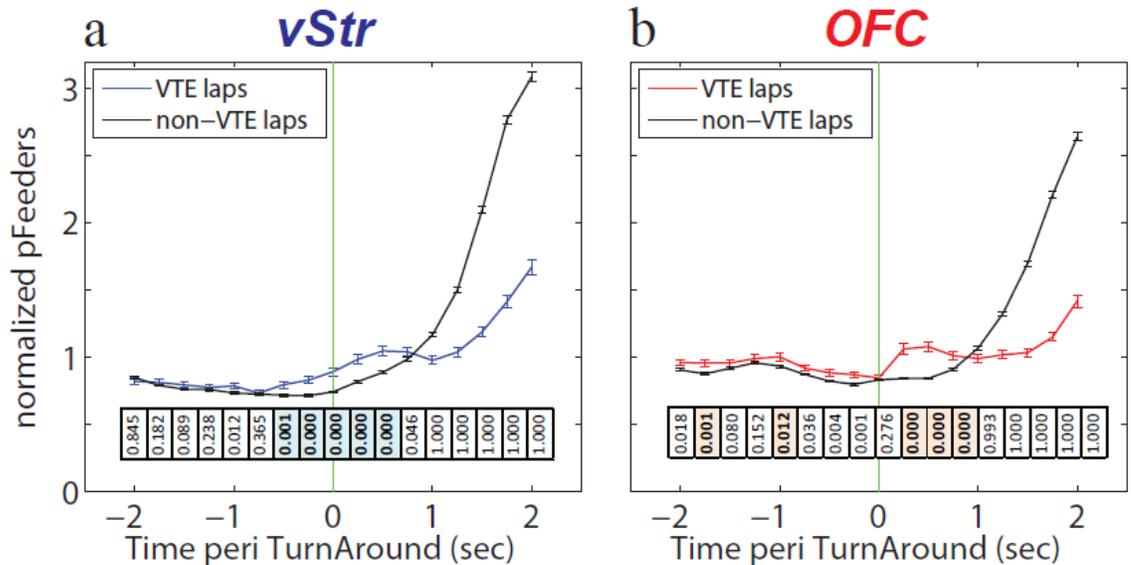


Figure 3-13. **Reward site decoding during VTE.** For each session with >5 cells in vStr and >5 cells in OFC (n=85 sessions), we calculated the average decoding for VTE passes and for non-VTE passes within-session. Lines represent the average across sessions, with error bars representing the standard error of the mean (SEM). (a) Feeder site decoding for vStr. VTE passes are aligned to TurnAround (see Methods and Figure S4) and non-VTE passes are aligned to the MidPoint of the choice point trajectory. VTE passes are shown in blue and non-VTE passes in black. Statistics show paired t-test (right-tailed) for each time bin, testing the hypothesis that VTE *pFeeders* was greater

than non-VTE *pFeeders*, with a Bonferonni correction for 17 time bins and two conditions (VTE vs. non-VTE) for  $\alpha=0.0015$ . P-values are shown directly below each time bin in both (a) and (b). Time bins where VTE was significantly greater than non-VTE are indicated by blue and pink shading, for vStr and OFC, respectively. **(b)** Feeder site decoding for OFC. VTE passes are shown in red and non-VTE passes in black. Statistical analysis as in (a). Decoding in OFC was significantly higher on VTE passes than on non-VTE passes, but only after the time of TurnAround, and after the time of significant decoding in vStr. Figure image from [144], reprinted with permission.

*PFeeders* values in OFC were also greater during VTE laps as compared to non-VTE laps (Figure 3-13 B), as seen in Steiner & Redish [120]. Notably, this increase in decoding occurred largely after the moment of TurnAround. Thus, during VTE, increased feeder site representations occurred earlier in vStr than in OFC, with the increase in vStr occurring before the moment of choice, and the increase in OFC occurring after the moment of choice. These differences in timing suggest that vStr precedes OFC in outcome valuation during deliberative behavior. This result could have implications for information processing in these two structures, as discussed below.

Covert representations of reward do not in themselves indicate which action the rat will ultimately take. In order to ascertain whether decoded neural signals were informative of the rats' choices, we examined the representation of each feeder site separately. Specifically, we hypothesized that if neural activity in either OFC or vStr was more predictive of what side the rat would ultimately choose, this should be reflected in differential representations of the chosen versus the unchosen feeder site. This was examined by plotting the subtraction at each time step of decoding to the chosen feeder site minus decoding to the unchosen feeder site (Figure 3-14). Values greater than zero denote stronger decoding to the feeder site that the rat will ultimately choose, indicating a neural representation of the rat's choice.

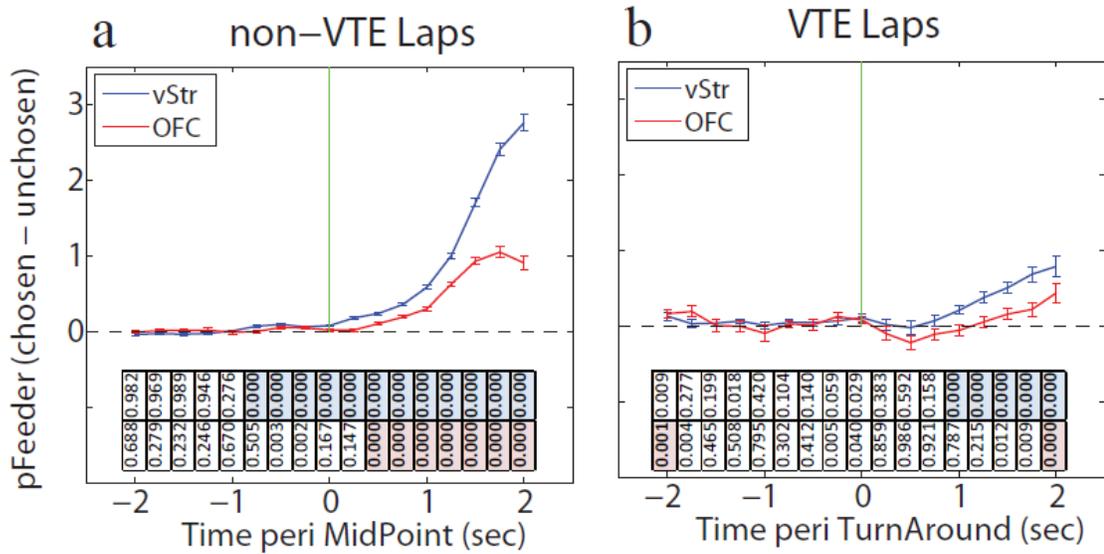


Figure 3-14. **Separate feeder site representations during choice.** Decoding values were averaged within-session, as in Figure 5. Lines represent the average across sessions, with error bars representing the standard error of the mean (SEM). **(a)** Decoding to the chosen feeder site minus decoding to the unchosen feeder site for non-VTE passes. The green line marks the MidPoint time. Values above the dotted line indicate greater decoding to the chosen feeder site. Statistics show t-tests for each time bin, with the null hypothesis that  $pFeeder$  was not significantly greater than zero (right-tailed), Bonferonni corrected for 17 time bins, two brain structures, and two behavioral conditions:  $\alpha=0.000735$ ). P-values are shown directly below each time bin in both (a) and (b). Time bins where  $pFeeder (chosen - unchosen)$  was significantly greater than zero are indicated by blue and pink shading, for vStr and OFC, respectively **(b)** Chosen side minus unchosen side decoding for VTE passes. The green line marks the TurnAround start time. Statistical analysis as in (a). Significant time bins for vStr preceded those for OFC on both non-VTE passes and VTE passes. Figure image from [144], reprinted with permission.

Looking at the magnitude and timecourse of  $pFeeder$  values ( $pFeeder$  here indicates decoding to a single feeder site), there were distinct differences between OFC and vStr, and between VTE and non-VTE laps. For both VTE laps and non-VTE laps, choice-predictive decoding become significantly greater than zero earlier in vStr than in OFC (Figure 3-14 A & B). This result

matches the results in Figure 3-13, because in both cases vStr preceded OFC, and vStr reached significance before time 0, whereas the OFC reached significance only afterward. However, this data does not completely align with the previous analysis. Unlike the result in Figure 3-13, for Figure 3-14 the temporal precedence of vStr is present for *both* VTE and non-VTE laps. Additionally, increased decoding to the chosen side reached significance earlier on *non-VTE laps* (for vStr and OFC), and the magnitude of the effect was greater for non-VTE laps (comparing absolute values between the two plots). Therefore, whereas covert reward coding showed increased VTE, choice-predictive activity was stronger during non-VTE laps.

However, these differences may not be surprising, given that the two analyses are measuring different phenomena. The analysis in Figure 3-13 is concerned with outcome evaluation. Increased reward site decoding selectively for VTE laps agrees with the role of VTE in deliberation and the prediction of outcomes [69]. Because VTE involves deliberation, one would predict that chosen value signals should emerge later during VTE events, as compared to “non-deliberative,” or habitual, passes through the choice point.

This reasoning explains why chosen value signals emerged earlier on non-VTE laps than on VTE laps. On non-VTE laps, which are more closely associated with alternation behavior (Figure 3-5), rats’ are likely employing a fixed strategy (i.e. “run to the opposite side”), and this earlier “commitment” in their behavior may be reflected in neural signals in vStr and OFC (Figure 3-14 A). Still, the temporal ordering was the same: vStr preceded OFC. Thus, task-relevant signals emerged earlier in vStr during both deliberative and non-deliberative behavior modes. This latter result matches a recent report in monkeys in which chosen value signals emerged earlier in vStr than in mPFC/OFC on a task that was overtrained and in which learning was explicitly minimized, and thus likely habitual in nature (Strait et al., 2015 [51], Figure 4C).

The analyses described above are spatial decoding analyses. Therefore, they could be sensitive to the particular trajectories taken by the rats. Because the actual paths towards the final choice point may proceed through different spatial locations, any spatial information that is present

in vStr or OFC could potentially generate the differences seen in Figures 3-13 and 3-14. Although previous decoding analyses on similar T-maze type tasks have not found significant spatial information encoded within vStr or OFC [120, 186], experiments by other groups have found spatial relationships within vStr and OFC on other tasks [38, 50]. To address this issue, control analyses were run to examine the contribution of spatial information to the decoding effects described above. Both spatial controls showed no difference in the magnitude or timecourse of decoding between OFC and vStr (Figures 3-15 & 3-16). Therefore, the differences between OFC and vStr seen in Figures 3-13 and 3-14 cannot be accounted for by possible differences in spatial tuning.

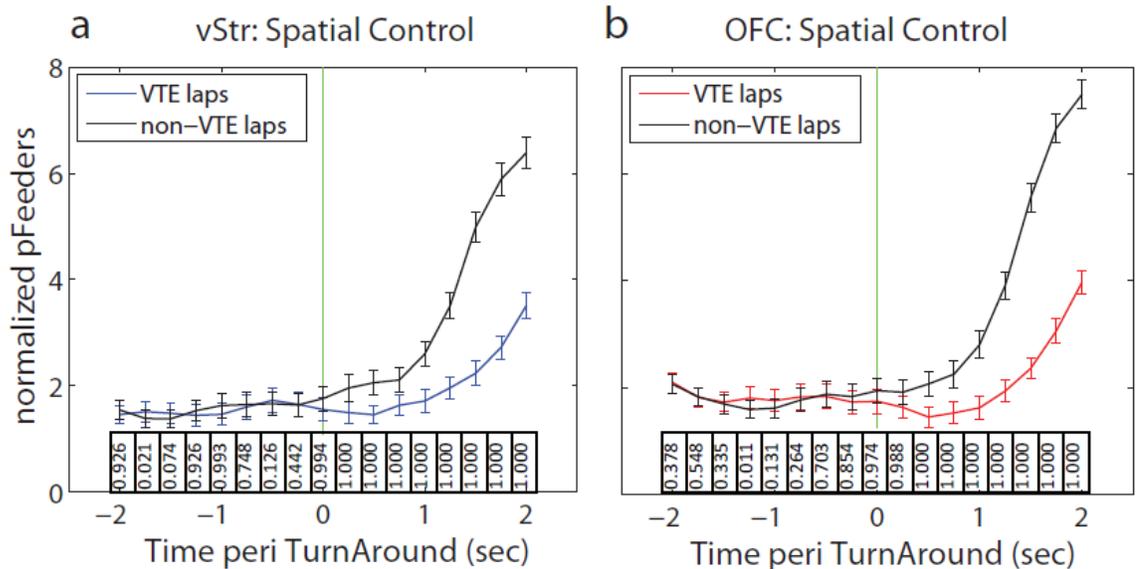


Figure 3-15. **Spatial control for Figure 3-13.** In order to control for potential spatial confounds, we recalculated the covert reward analyses using the expected firing rate of each cell given the spatial location of the animal. We first calculated the spatial tuning curve of each cell, and then substituted the average firing rate of the cell at the current location of the rat for the actual firing rate of the cell in the Bayesian decoding analysis. This removes any information not derived from the actual location of the rat and controls for any spatial differences in the location of the rat during the pass through the choice point. Data was averaged within session (n = 85 sessions). Lines represent the average across sessions, with error bars representing the standard error of the mean

(SEM). **(a)** Average feeder site decoding for vStr. VTE laps are aligned to the point of TurnAround and non-VTE laps are aligned to the MidPoint of the choice point trajectory. P-values are shown below each time bin as in Figure 3-11 (one-tailed t-tests, testing whether VTE greater than non-VTE). **(b)** Average feeder site decoding for OFC. Conventions as in (a). Note that for both vStr and OFC, the value of  $p_{Feeders}$  for VTE laps is even with or below that for non-VTE laps, indicating that the increases in VTE seen in Figure 3-11 are not due to spatial position. Figure image from [144], reprinted with permission.

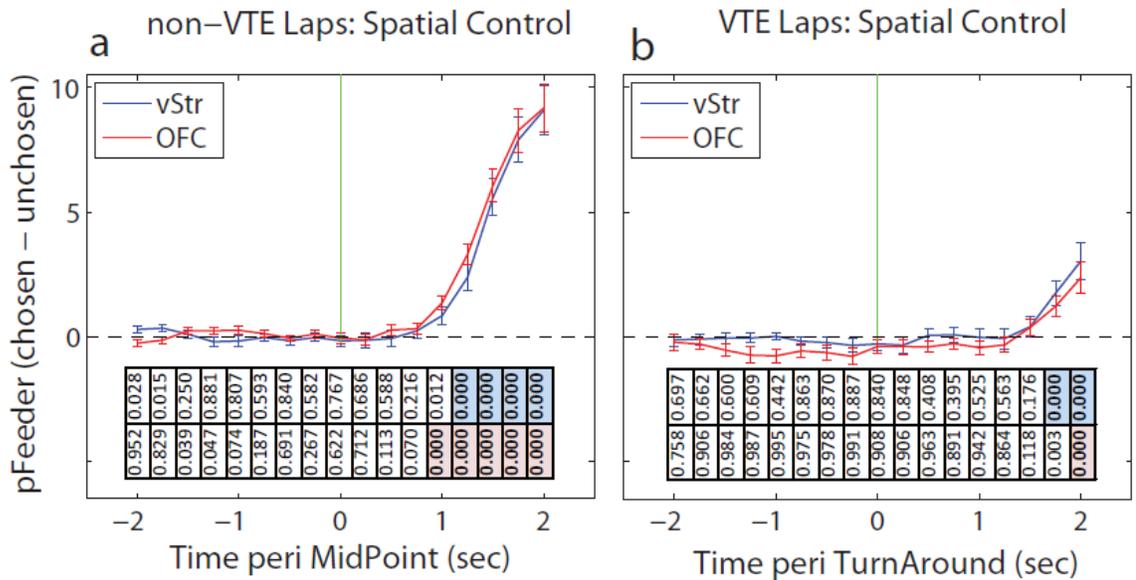


Figure 3-16. **Spatial control for Figure 3-14.** Same spatial control as above, but applied to the analysis from Figure 3-12. **(a)** Decoding to the chosen feeder site minus decoding to the unchosen feeder site for non-VTE laps. The green line marks the MidPoint time. Values above the dotted line indicate greater decoding to the chosen feeder site. P-values are shown directly below each time bin in both (a) and (b). Time bins where vStr or OFC was significantly greater than zero (one-tailed t-tests on each time bin—Bonferonni corrected— $\alpha = 0.000735$ ) are indicated by blue and pink shading, for vStr and OFC, respectively. **(b)** Chosen side minus unchosen side decoding for VTE laps only. The green line marks the TurnAround. For both structures, and both conditions (non-VTE & VTE), values were not greater than zero until after TurnAround. Importantly, the emergence of significance for vStr before OFC in Figure 3-14 is not seen here, indicating that the earlier timing seen in vStr in Figure 3-14 is not due to spatial position. Figure image from [144], reprinted with permission.

## Discussion

We made dual-structure recordings in orbitofrontal cortex and ventral striatum as rats performed an economic decision-making task, the spatial adjusting delay-discounting task. Behaviorally, rats on the DD task showed robust titration behavior. Over the course of the session, rats consistently moved the adjusting delay to a mean of around 5-6 seconds, and showed steady-state alternation behavior at the end of the session. This behavioral output provided a reliable measure of their indifference point, and showed that rats on this task were sensitive to the magnitude and delay to reward. By weighing magnitude against delay, rats showed that they were engaging in temporal discounting.

Rats also showed prominent vicarious trial and error behavior on the DD task. VTE was more frequent early in the session, during the titration phase, when the adjusting delay was changing, and less frequent later in the session, when the adjusting delay was relatively stable. This suggests that VTE behavior is engaged when some aspect or aspects of the task are dynamically changing. An analysis of the distribution of the two lap types (repeating versus alternation laps) in relation to VTE events showed that when VTE occurred, rats choose each maze arm with equal likelihood. This provides strong evidence that VTE reflects a deliberation on the part of the rat. Non-VTE laps co-occurred selectively with alternation laps, suggesting that non-VTE passes through the choice point took place at a time when the rats' were engaged in a fixed strategy, indicative of habitual behavior. The neural signatures of deliberation during VTE were observed both early and late in the session<sup>11</sup>, supporting the idea that VTE and non-VTE events reflect two distinct cognitive states.

The OFC and vStr are centrally involved in delay-discounting behavior [187]. Additionally, both of these structures exhibit neural signatures of reward-related processing during VTE behavior [48, 120]. In the case of delay-discounting, others have found a great deal of overlap

---

<sup>11</sup> Figure 3-11 & Figure 3-12 were similar when plotted using only the first 50 or using only the last 50 laps.

in the functional activity seen in OFC and vStr [12, 14]. In the case of VTE, similar signals have also been reported, but using different methodologies. Simultaneous recordings from OFC and vStr on the DD task allowed me to address both of these questions, and additionally, to examine the neural signatures of VTE on an economic decision-making task, which had not yet been addressed.

For the delay-discounting component, we saw similar activity profiles in OFC and vStr neurons. In particular, both structures showed similar percentages of reward-responsive cells, and roughly similar degrees of magnitude and delay coding, with a “mixed” coding strategy apparent in both structures. However, we did find significant differences between these brain areas in terms of the timing of outcome-related and choice-related neural activity. The ventral striatum exhibited a covert representation of reward before the moment of choice during deliberative behavioral modes, similar to that seen by van der Meer & Redish [48]. We did not observe a similar pre-choice increase for OFC. Instead, the data presented here, and that from a previous study [120], found that reward site representations increased in OFC *after* the rat had committed to its decision. In line with this result, the vStr also preceded OFC in distinguishing the two feeder sites before the animal made its choice. These results imply that vStr and OFC are engaged at different times during decision-making: vStr before the choice is made, and OFC after the choice is made. This dissociation in timing has implications for theories of orbitofrontal and ventral striatal function in decision-making.

The pattern of behavior that we observed on the DD task provided the opportunity to compare flexible with non-flexible decision-making, and simultaneous recordings allowed for a direct comparison of OFC and vStr activity. Comparing VTE to non-VTE laps, we observed an increase in the feeder site representation in vStr before the moment of choice. This increase in feeder site representation during VTE supports the assertion that the vStr provides a covert reward signal selectively during flexible decision-making [48].

In contrast to the timing of the covert reward signal in vStr, there was no increase in the feeder site representation in OFC before the moment of choice. The absence of a covert reward

signal in OFC before the rat selected his action indicates that the OFC does not necessarily show planning signals of the kind seen in vStr. Although we did not see covert expectation of reward in OFC, the study by Steiner & Redish [120] did report an increase in the feeder site representation in OFC during VTE. In that study, the increase occurred *after* the turnaround event, and thus *after* the rat had changed direction. Such a signal may represent information about the rat's choice after it has been taken; for example, an expectation of reward [120], a representation of the state the animal is in [45], or a linkage between the chosen action and the eventual outcome [188].

Covert reward site representations during VTE were present before the moment of choice in ventral striatum and potentially after the moment of choice in orbitofrontal cortex. This suggested that information about the rat's impending action might also be differentially expressed in OFC and vStr. The relative contributions of OFC and vStr to action selection are poorly understood, and the contribution of OFC in particular to action selection is controversial [20, 35, 36, 189]. In cued-based economic tasks, conditioned stimuli indicate the relative value of the choice options at the time(s) of cue presentation. Because neurons in OFC signal the value expected rewards [17, 53], and because animals are very likely to choose the higher-valued reward (although not always), OFC neural activity during cue presentation is highly indicative of which action the animal (or human) will ultimately take. Relatively little is known about how choice-predictive information is represented in OFC and vStr in the absence of explicit sensory cues. The spatial delay discounting task presented no cues before the animal made its choice, and therefore presented an opportunity to investigate choice-related neural activity in the context of internally generated, self-initiated decisions.

Comparing neural activity in OFC and vStr side by side during decision-making, we found that neural representations of reward and choice emerged earlier in ventral striatum than in orbitofrontal cortex. Non-local reward representations in vStr preceded those in OFC specifically during deliberative VTE behavior. Covert signals informative of the animal's impending choice also emerged earlier in vStr. Taken together, these results suggest that expectancy signals of key

decision variables are encoded by ventral striatum. Signals in OFC emerged after the rat has made its decision, possibly encoding information about the value of the chosen action. This data emphasizes the importance of vStr activity in planning actions during deliberative behavior and argues for the incorporation of vStr into network models of deliberative decision-making.

# Chapter 4: Local Field Potential Analyses

## Introduction

Decoding analyses from simultaneous ensemble recordings in OFC and vStr demonstrated that the vStr preceded the OFC in showing reward-related representations during VTE, and that these reward-related representations occurred before the moment of choice in vStr and after the moment of choice in OFC. Additionally, the vStr also preceded the OFC in manifesting choice-predictive activity. These results have implications for how vStr and OFC may be processing information during the decision-making process.

As put forward in Chapter 3, these results suggest that the vStr is involved in a planning or “look-ahead” process during VTE, generating an expectation of reward that can be used to guide the animal’s decision. Although this role is more commonly attributed to the OFC [57], the OFC was active *after* the moment of choice on the DD task, suggesting a different role in the deliberative process. In particular, OFC activation after the point of turnaround may signal information about the chosen action, which is necessary for proper credit assignment [188].

The relative timing of decoded activity in OFC and vStr raises an intriguing possibility; namely, that the vStr is sending information to the OFC during the decision-making process, influencing signaling in the orbitofrontal cortex. A useful tool for addressing questions of this nature is through local field potential (LFP) analysis.

Local field potential signals have proved essential for studying interactions between brain structures [130]. The local field potential measures aggregate neural activity over a larger spatial scale (on the order of hundreds of microns), reflecting the activity of a large population of neurons, serving as a complement to single unit recordings. Furthermore, most circuits in the nervous system are predisposed to oscillate at specific frequencies [190]. These oscillations, readily measured in the LFP, are thought to mediate information processing in a number of contexts [191]. A large body

of literature shows that long-range synchrony between brain areas correlates with numerous behavioral and cognitive events, including successful memory encoding [192], enhanced visual attention [193], spatial working memory [194, 195], and anxiety states [196]. Because of the importance of oscillatory activity in inter-regional communication [197], LFP recordings can provide valuable insight into the functional interactions between brain structures.

One purported mechanism by which synchronous activity is thought to influence behavior is through cell assembly dynamics [198]. Peaks and troughs in the local field potential correspond to windows of rhythmic inhibition and excitation, particularly at gamma frequency [199]. Synchronous discharge in projection neurons results in temporal summation downstream and more effective information transfer [200]. Differences in the efficacy of information transfer could affect behavior on a moment to moment basis. Additionally, synchronous activity could enable learning through synaptic plasticity mechanisms [201]. Co-active cell assemblies in different brain areas could therefore be selectively strengthened (or weakened) between periods of inhibition via Hebbian plasticity mechanisms [202, 203]. Changes in LFP power at specific times could act as a gain control for these effects, either in terms of information transmission [204], or in terms of synaptic strength changes thought to be responsible for learning [205].

Another mechanism by which synchrony could influence information processing is through the segregation of different inputs. Just as synchrony between structures can allow for efficient communication between them, desynchronization or a change in frequency could block certain inputs, allowing a given structure to switch between different inputs streams [206]. An example of this can be found in the hippocampus. Fast gamma oscillations in area CA1 were synchronized with fast oscillations in the medial entorhinal cortex (MEC), while slow gamma oscillations in CA1 were coherent with slow gamma oscillations in area CA3 [207]. These gamma sub-bands in CA1 occurred largely on different theta cycles, possibly allowing for the selective routing of information from either MEC or CA3. Thus, changes in oscillatory power, as measured

by the local field potential, could reflect temporal segregation of information from different sources.

In order to examine the dynamics of information transmission between the OFC and the vStr, as well as the hippocampus (HIPPO), local field potential data from the same dataset of DD recording sessions were analyzed. Although the OFC and vStr are both critically involved in value-based decision making, interactions between these structures has not been well studied. Simultaneous recordings in our dataset allowed us to investigate the relationship between these two structures. If the OFC and the vStr are functionally coupled during value-based decision-making, then this should be apparent in measures of co-activity, like coherence. If the vStr is leading the OFC, as suggested by the timing data in Chapter 3, then this should be evident in directional measures like Granger causality (GC). For a description of the methods used for LFP analysis, see Chapter 2: Methods.

## **Results**

### Fundamental frequencies in the OFC and vStr

Interactions between brain structures occur at specific frequencies [190]. Prior work has identified gamma oscillations as being important in both the vStr [128] and the OFC [129]. What are the dominant frequencies in each structure on the delay-discounting task? Figure 4-1 shows the power of the LFP signal as a function of frequency, in “power spectral density” (PSD) plots, separated by brain region. Power in the extracellular local field potential falls off with a roughly  $1/f$  relationship to frequency. Significant peaks along this line indicate frequencies with increased power. The OFC shows a peak at 50 Hz (Figure 4-1 A), the vStr shows a sharp peak at 50 Hz and a broad shoulder centered at 80 Hz (Figure 4-1 B). These two frequency bands will be referred to as “gamma50” (45-55 Hz) and “gamma80” (70-85 Hz). LFPs from the hippocampus show a distinct peak at theta frequency (5-12 Hz) (Figure 4-1 C), as would be expected from a recording site near

the hippocampal fissure [208]. Thus, the PSDs indicate that oscillations in the gamma range in OFC and vStr, and in the theta range in hippocampus are spectral “regions of interest.”

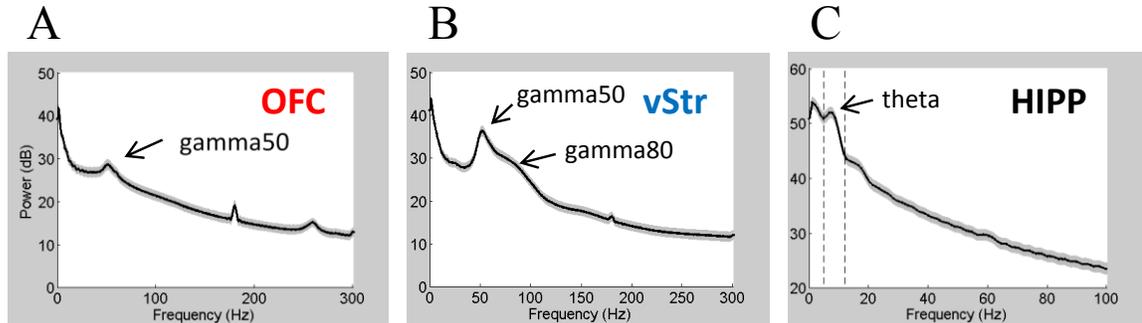


Figure 4-1. **Spectral peaks in recorded areas.** Power spectral density plots of average power  $\pm$  SEM as a function of frequency in OFC (A), vStr (B), and hippocampus (C).  $n = 140$  sessions for OFC and vStr, and 89 sessions for hippocampus. Both OFC and vStr have prominent gamma peaks at 50 Hz (“gamma50”). Vstr has a broad should of gamma activity between 70 and 100 Hz (“gamma80”). Recordings from the hippocampal fissure show a peak at  $\sim 8$  Hz, the theta rhythm. Dashed lines are at 5 Hz and 12 Hz.

This assessment is corroborated by looking at the raw data. Figure 4-2 shows a screenshot from the Neuralynx display taken prior to the start of the task, for animal R244. Brackets indicate the vStr channels and the OFC channels, and the hippocampus channel is indicated by the arrow. Strong gamma oscillations are readily apparent in the vStr channels. They can occur spontaneously, as in this sample, but as will be seen later, they are modulated by task-dependent variables. It is also clear by visual inspection that gamma oscillatory activity occurs in discrete bouts, as will be discussed below, and as has been reported previously [48].

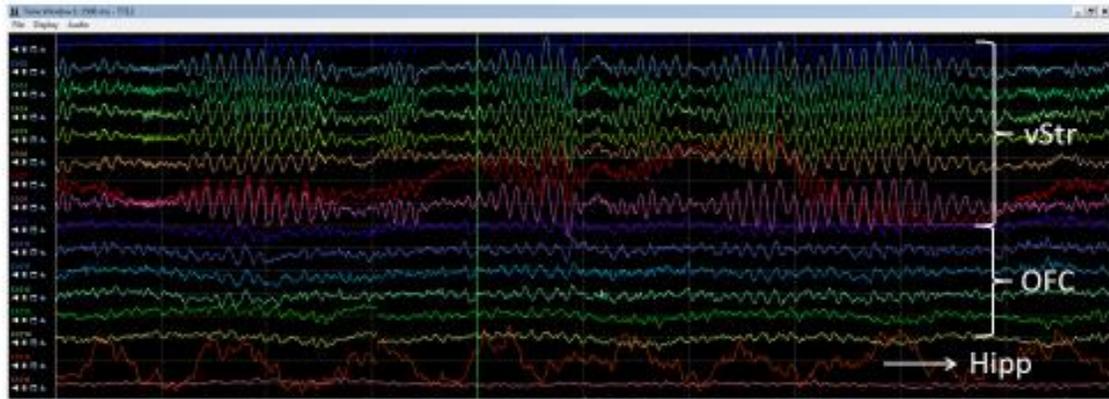


Figure 4-2. **Example oscillatory activity in the recorded brain areas.** A screenshot from R244-2012-08-24 prior to task start. Note the prominent gamma oscillations in the vStr channels, and also in the OFC channels at lower amplitude. The auxiliary electrode in the hippocampus shows a strong theta oscillation. Duration of the screenshot is 1000ms.

Two other features are worth noting in the screenshot. The OFC channels show gamma oscillations, although the amplitude is smaller than in vStr. Gamma bouts in OFC also appear to co-occur—at least some of the time—with the vStr gamma bouts, suggesting there may be coherent oscillations at this frequency. Lastly, the hippocampal channel shows a good example of theta activity. In this example, one can count eight cycles within the one second span of the sample (8 Hz rhythm). Although this screenshot shows “good” LFP data, these characteristics were common in my recordings. Gamma oscillations are prevalent in the vStr [128], and they increase in power in the striatum along a dorsolateral to ventromedial gradient [148]. Gamma oscillations in the rodent orbitofrontal cortex have also been reported, with neural rhythmicity in this frequency band playing a role in associative learning and decision-making [150, 209].

We also identified peaks in the power spectra by calculating the power correlation between frequencies [159]. This technique is advantageous in that it does not involve any *a priori* filtering, it is a more sensitive measure for non-stationary data, and it reveals interactions between frequencies. Figure 4-3 displays the cross-correlation of power at different frequencies—termed “self-coherence” [159]—for OFC, vStr, and HIPP. These data show “fundamental” frequencies at

50 Hz for both OFC and vStr, and at theta frequency in the hippocampus. In addition, both OFC and vStr show detectable peaks in the beta range (~25 Hz), and these peaks are correlated with gamma50 power (see arrows that indicate the off-axis, symmetrical peaks at 25 and 50 Hz). Beta band activity has been linked to task-related sensorimotor modulation in the dorsal striatum [210, 211] and to habit learning in the ventromedial striatum [212]. The relationship between beta band and gamma band oscillations has not been investigated in the vStr or OFC.

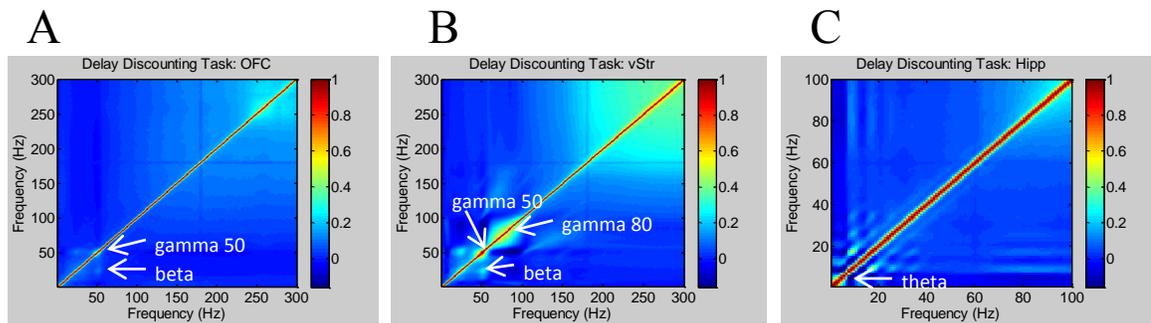


Figure 4-3. **Self-coherence plots for OFC, vStr, and HIPP.** Each plot shows correlation coefficients between power at different frequencies for OFC (A), vStr (B), and hippocampus (C). Warmer colors indicate higher values and positive correlations. Dark colors (below zero) indicate anti-correlations. Areas of high power show up as “blobs” about the diagonal, because oscillations in biological signals tend to co-occur over a range of frequencies (i.e. gamma50 does not occur strictly at 50 Hz). High (or low) off-axis areas indicate power correlations between frequencies. In OFC and vStr, gamma50 tends to co-occur with beta activity.

In Figure 4-3 B, the vStr shows an area of high correlation in the gamma80 range, spanning approximately 70-100 Hz, consistent with the PSD (Figure 4-1 B). Interestingly, there is a marked anti-correlation between the gamma50 and gamma80 bands, indicated by the black stripes with negative correlation values. This suggests that in the vStr, gamma50 and gamma80 bouts occur at different times, as reported previously [119]. This effect is illustrated in the example spectrograms shown in Figure 4-4. The temporal segregation of gamma50 and gamma80 bouts can be seen in vStr (Figure 4-4 B), but also, to a lesser extent, in OFC (Figure 4-4 A). These data suggest that

power in the 50 Hz and 80 Hz range are temporally distinct frequency bands in the vStr, which may serve different functional roles [48]. Taken together, the PSD plots and the self-coherence plots show that the OFC and vStr local field potentials are characterized by fundamental frequencies at 50 Hz, show correlations with beta power at 25 Hz, and additionally, the vStr contains a significant band of high power in the 80 Hz range, anti-correlated with gamma50 power.

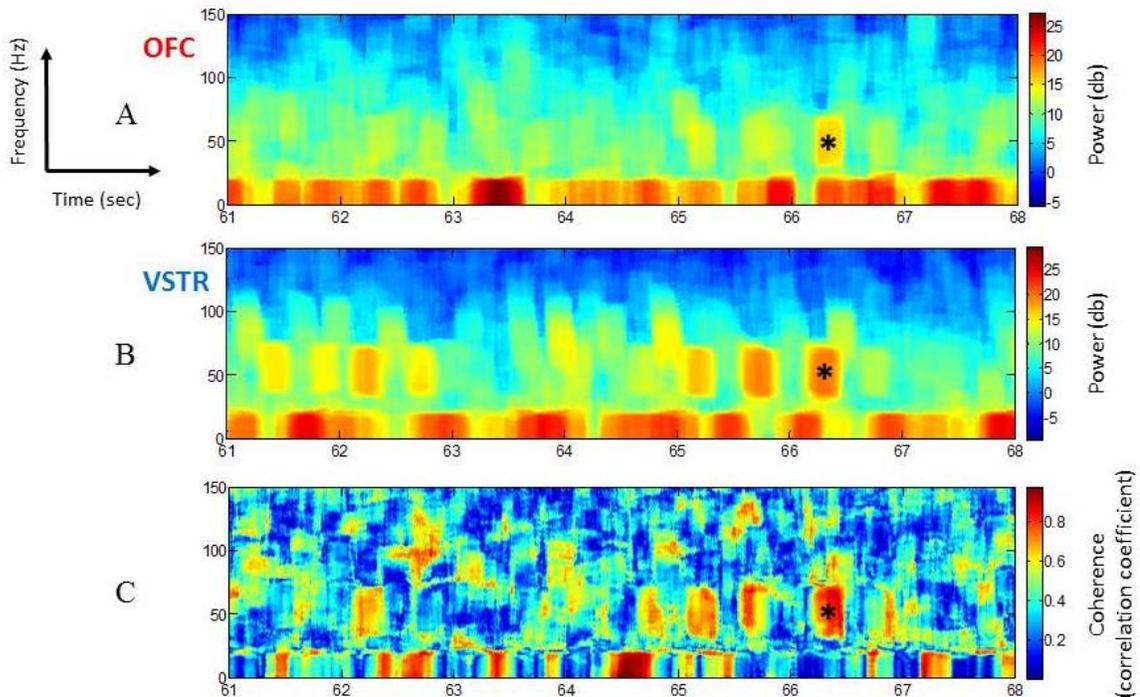


Figure 4-4. **Example power-frequency spectrograms and coherogram between OFC and vStr.** Panels (A) and (B) show spectrograms for a short time window during behavior on the DD task, for OFC, and vStr, respectively. Session R226-2011-08-19. Note the alternating bouts of gamma50 and gamma80 power in the vStr. Panel (C) shows coherence between the two time series in A and B. Warmer colors indicate greater coherence. Note the moments of high coherence, especially in the gamma50 range (example shown with black asterisks).

Gamma50 and gamma80 oscillations are differentially modulated by task events

The data above supports the interpretation that gamma50 and gamma80 oscillations in the striatum are temporally distinct [128]. In order to examine the relationship between gamma

oscillations and task-relevant events, PETHs of gamma50 and gamma80 power were constructed for the choice point and reward receipt epochs.

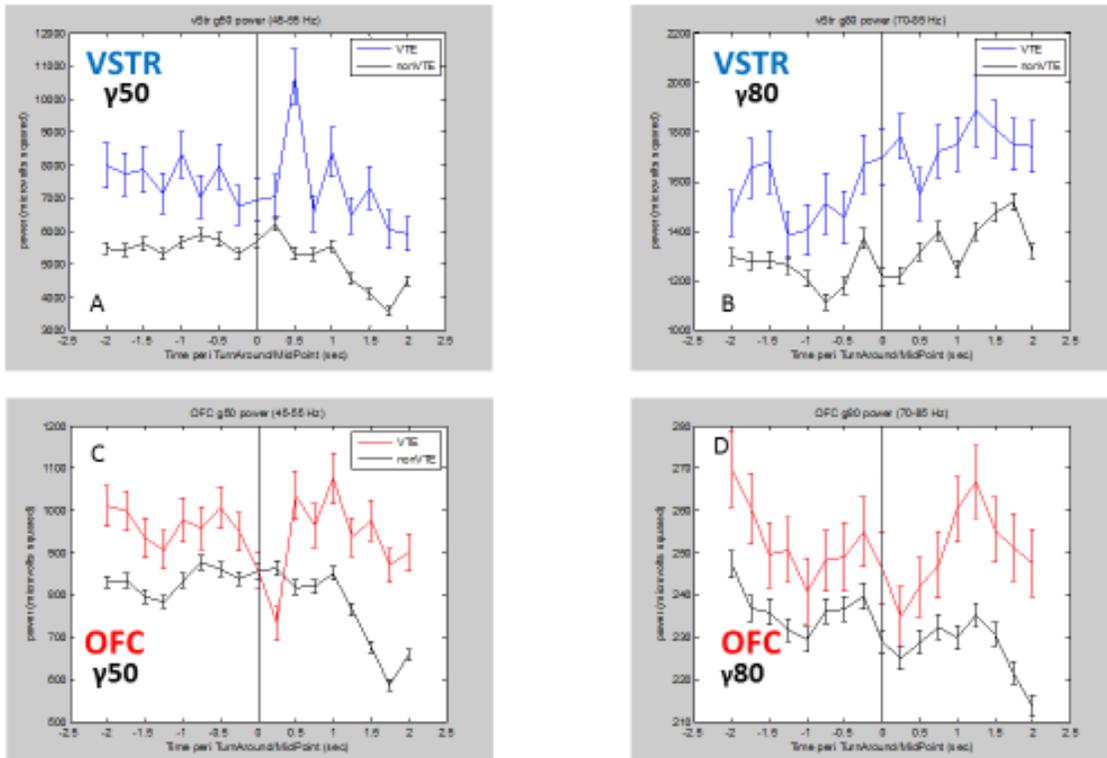


Figure 4-5. **Gamma power at the choice point.** Gamma power PETHs aligned to the TurnAround Point (VTE laps, blue/red) or the MidPoint (non-VTE laps, black) for vStr (A) & (B) and OFC (C) & (D). The vStr shows an increase in gamma50 power after the point of TurnAround, selectively on VTE laps.

At the choice point, gamma50 power in vStr increased transiently after the point of TurnAround (Figure 4-5 A), while gamma80 in the vStr and gamma power in the OFC did not show a clear change at the time of decision (Figure 4-5). This increase in vStr gamma50 power matches the increase in gamma50 seen in van der Meer & Redish [119] at the time of reversal (when the rat reverses direction on “wrong way” laps), and is most likely related to the increase in gamma50 seen throughout the striatum prior to movement onset [152]. This increase was not observed in OFC (Figure 4-5 C). Rather, OFC gamma50 appeared to decrease at the time of

TurnAround. Overall, gamma power was somewhat noisy at the choice point. Interestingly, gamma power was higher during VTE laps than non-VTE laps for all conditions. This could indicate greater within-structure synchronization, as increased gamma band power has been linked to increased spike-field coherence during visual attention [213, 214].

Around the time of reward receipt, gamma50 power in vStr and OFC decreased transiently just before the feeder fire event and then rose sharply, showing a “double bump”, likely corresponding first to a response to the cue and then to the arrival of the food reward ~ 2s later (indicated by black and red lines, respectively) (Figure 4-6, A & B). Gamma50 power was elevated and sustained during the adjusting delay (Figure 4-6, A & B). Gamma80 power dropped steeply after the feeder fire event, before returning to baseline (Figure 4-6, C & D).

These dynamics are similar to those observed in the ventral striatum by van der Meer & Redish [119]. In those data, gamma50 power showed a peak at the time of arrival at the feeder site and another peak and sustained elevation thereafter (Figure 9 from [119]). Donnelly et al. [215] likewise found that gamma50 power in mPFC and NAc showed transient increases during reward-related cue presentation and during reward consumption. In contrast, in our data and in that of van der Meer & Redish [119], gamma80 power decreased sharply and transiently at the time of reward receipt.

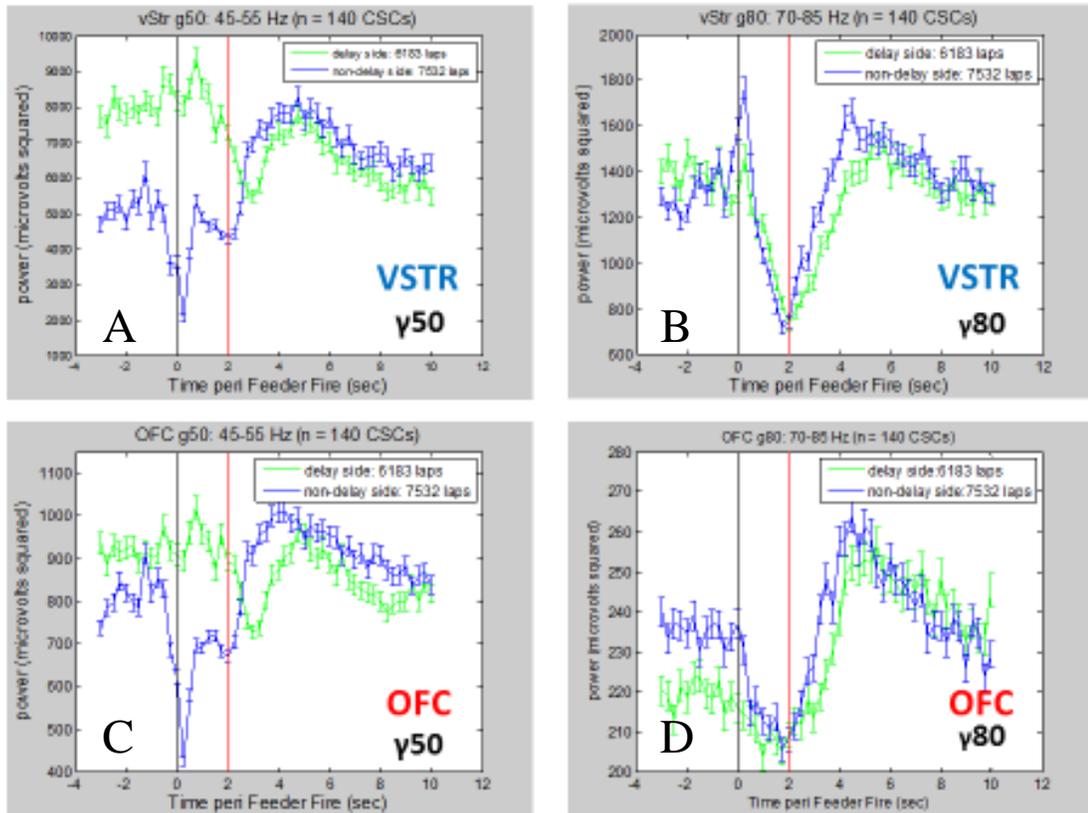


Figure 4-6. **Gamma power around the time of reward receipt.** Gamma power PETHs aligned to the time of Feeder Fire. Black lines indicate the time of Feeder Fire and red lines indicate the approximate time of reward receipt. Gamma50 power shows sustained, high activity during the delay (green lines), and a phasic decrease and then a “double bump” after the feeder fire cue and the arrival of the reward—(A) and (C). Gamma80 power shows a phasic decrease after the feeder fire cue on both lap types—(B) and (D).

Interestingly, in the data by van der Meer & Redish [119], gamma50 and gamma80 power remained relatively flat during error laps in which the rat did not receive reward after arrival at the feeder site (Figure 9a from that paper). The delay-discounting task does not permit the comparison of error versus correct laps, because by design, there is no incorrect choice. The animals received reward on every trial (except for a rare handful of trials in which the animals skipped the feeder). However, because reward magnitude differed on the DD task (unlike the Multiple-T maze), the response to different reward amounts can be compared. On the DD task, gamma50 power in the

vStr especially was higher before the time of the feeder trigger event for the 3 pellet reward, as compared to the 1 pellet reward (Figure 4-6 A), suggesting that gamma50 power in the vStr differentiates *anticipated* reward magnitude.

After the time of food pellet arrival (indicated by the red line), the magnitude and shape of gamma50 is similar for both small and large reward receipt, albeit a little sluggish on the delay side laps. This later activity is perhaps due to a small amount of uncertainty on the rats' part as to when the food will arrive, as compared to the shorter time between successive events on the non-delay side. This data indicates that gamma50 power does not distinguish between large and small rewards *at the time of reward receipt*. The dynamics of gamma80 power for both vStr and OFC were similar for large and small reward events, showing no evidence of modulation by reward magnitude.

By far the most striking feature of these data are that gamma50 and gamma80 show nearly opposite patterns of activity. Gamma50 shoots up and gamma80 drops precipitously after the feeder fire event. These data very closely match the dynamics seen in [119]. They are also partially consistent with the study by Kalenscher et al. [149], in which all of the LFPs with significant reward-related gamma power had higher activity at gamma80 before arrival at the feeder site than after.

These data may also be consistent with another report of a “switch” in gamma power seen around the time of reward receipt [148] (Figure 4 of that paper). On that task, arrival at the feeder ports was coincident with an interruption in gamma50 power and a transient burst of gamma80 power. This too can be seen in our data if we look at the time of feeder fire (which for non-delay side laps, approximately matches the time of feeder arrival). Although the change is fast, one can see that gamma50 power drops sharply just before feeder fire (before the “double peak”) and at the same time that gamma80 power shows a transient increase at that time (before the larger, phasic decrease). The timing of these effects was faster than the “switch” between 50 and 80 Hz in the Berke paper [148]. So it is unclear if the exact timing of these two analyses matches up. However,

the consistent factor among each of these reports is that gamma50 and gamma80 power appear to alternate around the time of reward receipt.

#### Coherence between OFC and vStr on the delay-discounting task

If the OFC and vStr are interacting on the DD task, then they should show coherence in the time-frequency domain (for a description of how coherence was calculated, see Chapter 2: Methods). Figure 4-4C showed an example trace with bouts of coherent activity between OFC and vStr at different frequencies, most notably around 50 Hz. In order to systematically evaluate areas of coherence between OFC and vStr, coherence spectra were obtained from paired OFC-vStr LFP time series for each session, with a sliding window of 1 second and 50% overlap. Data were averaged for each session. The across-session average coherence is shown for each animal in Figure 4-7 A-E, and the average taken over all sessions for all animals is shown in Figure 4-7 F. These data show clear peaks in the coherence spectra at delta (4 Hz), beta (~25Hz), and gamma50 (~55 Hz) frequencies, and a broad “shoulder” at gamma80 (70-100 Hz). Thus, the OFC and vStr show strong and frequency-specific bands of coherence on the delay-discounting task, suggesting a possible functional interaction between the two areas.

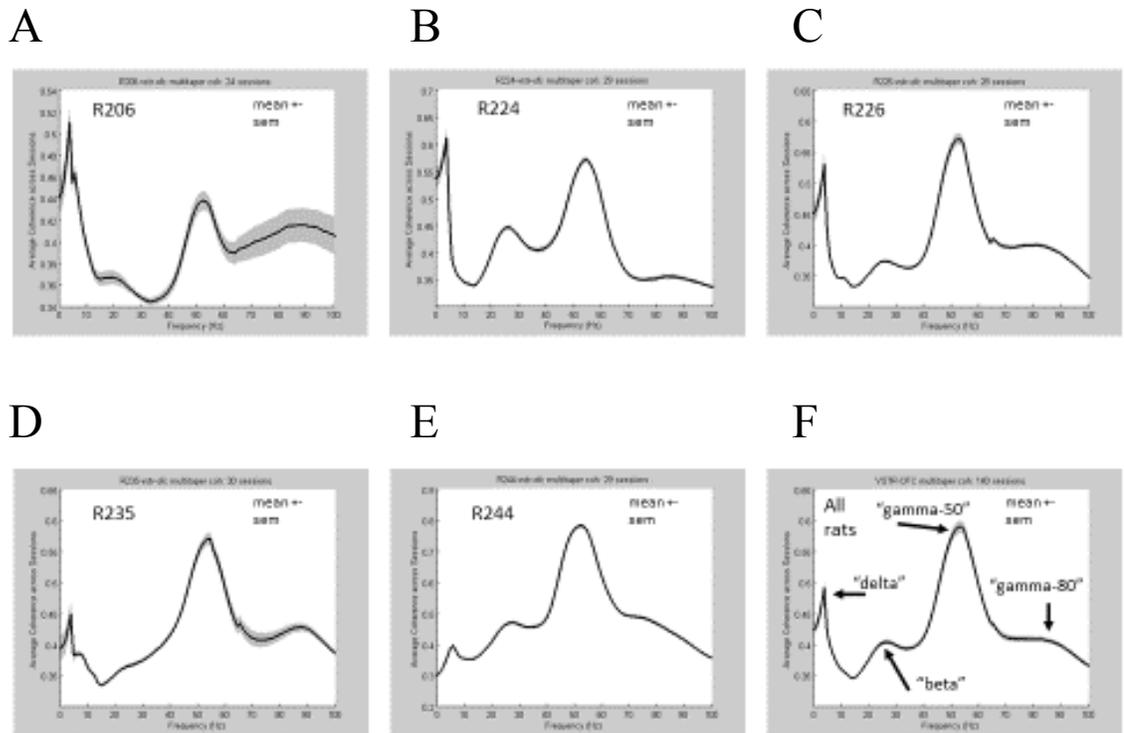


Figure 4-7. **Average coherence spectra across animals.** Each plot shows the mean coherence spectrum  $\pm$  SEM between simultaneous OFC and vStr recordings. Data are averaged across sessions, for each rat used in the LFP analyses (A-E). The bottom right hand plot (F) shows the average across all sessions taken together (i.e. average over all subjects:  $n = 140$  sessions). There are distinct peaks in OFC-vStr coherence at delta, beta, gamma50, and gamma80 frequencies.

To investigate synchronization between the OFC and vStr during task-relevant epochs, PETHs were constructed for coherence in the gamma50 and gamma80 frequency bands, as they were for gamma power (above). During choice point passes, there was a “bump” in gamma50 coherence after the moment of TurnAround, selectively for VTE laps. This increase was not present in the gamma80 range (Figure 4-8 A vs. B). This increase in gamma50 coherence matches the increase in power seen in Figure 4-5. Whether this increase in power and coherence at gamma50 is due to the decision-making process or simply due to movement initiation [152] is undetermined with this analysis. Further analyses might be able resolve this issue.

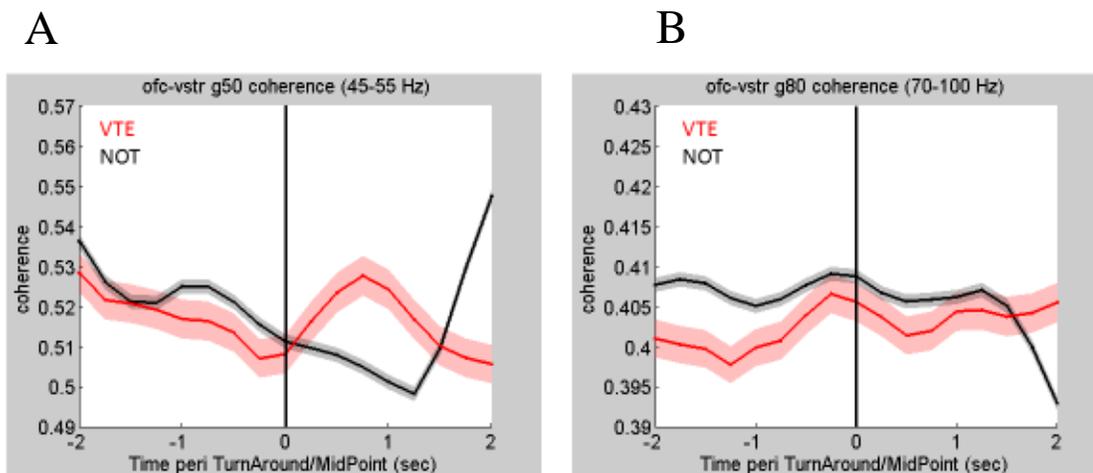


Figure 4-8. **OFC-vStr coherence at the choice point.** Each plot shows the mean coherence  $\pm$  SEM for gamma50 (**A**) and gamma80 (**B**), aligned to the time of TurnAround, or the MidPoint, for VTE laps (red) and non-VTE laps (black), respectively. Note the increase in gamma50 coherence after the time of TurnAround on VTE laps.

Coherence at gamma50 and gamma80 aligned to Feeder Zone Entry (Figure 4-9) showed similar dynamics to the average power in OFC and vStr at these times (Figure 4-6). Gamma50 coherence showed a “double bump” after entry into the feeder zone, likely corresponding to the auditory cue at zone entry and subsequent food delivery. Gamma80 coherence showed a phasic decrease. In both cases, VTE and non-VTE laps were highly similar. The alternating dynamics in the two gamma bands suggests that vStr and OFC become more strongly coupled at 50 Hz during reward anticipation (and receipt), and become less coupled in the 80 Hz range during this time.

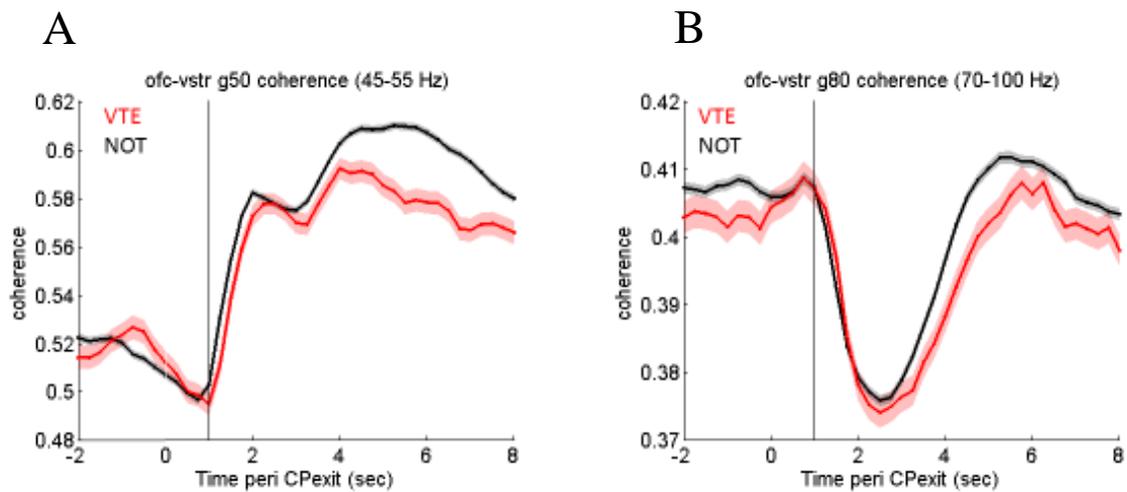


Figure 4-9. **OFC-vStr coherence at the time of reward receipt.** Each plot shows the mean coherence  $\pm$  SEM for aligned to Choice Point Exit (equivalent to feeder zone entry) for (A) gamma50 and (B) gamma80.

#### Directional influences between OFC and vStr on the delay-discounting task

The coherence data above suggest that OFC and vStr undergo closely matched fluctuations in power over time. However, coherence is not a directional measure. It does not differentiate between signal A leading signal B, or vice versa. Directional measures, on the other hand, can estimate the directionality and strength of information transmission between two simultaneously recorded signals [216, 217]. The results from Chapter 3 indicated that the vStr led the OFC in representing value-related and choice-predictive information during the decision-making process. The results presented in this chapter indicate possible functional coupling between OFC and vStr as measured by coherence. In order to determine whether the vStr directly influenced OFC activity, two directional measures were applied here to estimate directed functional connectivity between LFP signals in OFC and vStr; (1) Granger causality [162] and (2) a recently developed cross-correlation technique [171].

By definition, a time series A “Granger causes” time series B if the information in the past of time series A helps predict the future values of time series B, above and beyond the predictive

power of past values of time series B [218]. Granger causality has proven to be a powerful analytical tool for making inferences about causality between the sources of simultaneously sampled time series [164]. Here, we apply G-causality analysis to pairs of continuously sampled LFPs (one electrode in OFC, one electrode in vStr), but in principle, it could be applied to an arbitrary number of simultaneously recorded signals. All G-causality analyses presented here were made with custom-written MATLAB code adapted from the Multivariate Granger causality toolbox (MVGC) [162], freely available online. For a description of the methods used to calculate G-causality and a link to the toolbox, see Chapter 2: Methods.

To get a sense of the relationship between OFC and vStr neural signals at a macroscopic level, we applied G-causality analysis to paired LFP signals at the whole session level. This analysis yields a pair of values for each session, the “time-domain” Granger causality from OFC to vStr, and from vStr to OFC. These data are plotted in Figure 4-10 A. The x-axis represents GC values for OFC→vStr and the y-axis represents GC values for vStr→OFC. Values that fall above the identity line indicate sessions with stronger granger causality in the vStr→OFC direction, and those that fall below the identity line indicate sessions with stronger Granger causality in the OFC→vStr direction. Out of 140 sessions, 96 fall above the identity line (44 below). This difference was highly significant ( $p < 10^{-4}$ , Wilcoxon signed rank test). Therefore, at the session-wide level, the vStr is more often driving the OFC than vice versa.

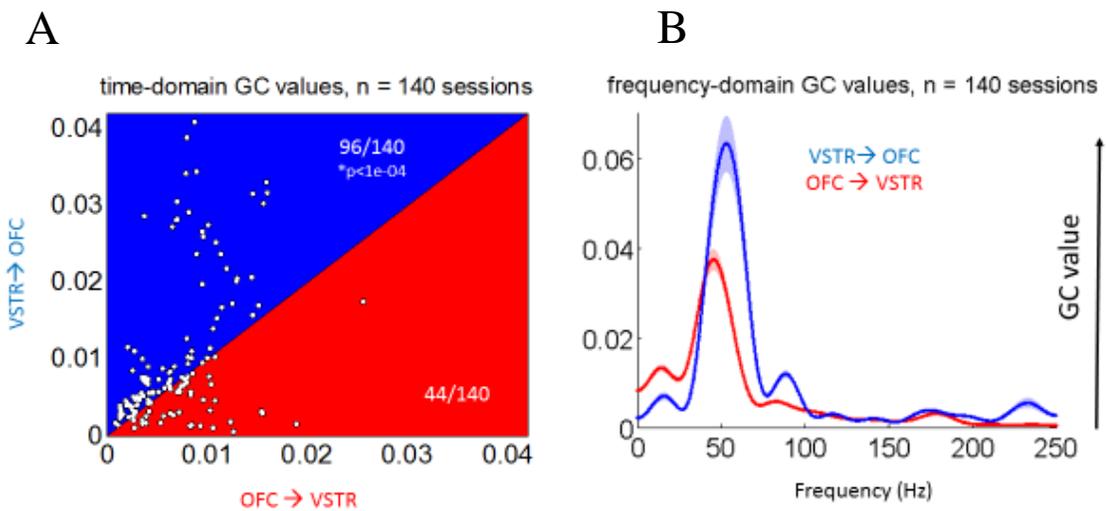


Figure 4-10. **Average granger causality between OFC and vStr.** Panel (A) shows time-domain GC value pairs (OFC→VSTR & VSTR→OFC) for all sessions. Each point represents the average for both values over an entire session. The red area marks greater GC values from OFC to vStr, and the blue area marks greater GC values from vStr to OFC. Panel (B) shows the same data decomposed into its spectral components. Each session yields a vector of values at each frequency for both the vStr→OFC and OFC→vStr directions. Data are averaged across sessions,  $\pm$  SEM. \*x-y significantly different from zero.

Time-domain granger causality can be decomposed into the relative contribution of its different frequency components [165, 166]. Under this formulation, time-domain granger causality is equal to the sum of the spectral components over all frequencies, from zero to the Nyquist frequency [163]. The spectral “factorization” of the time-domain GC values is shown in Figure 4-10 B. This plot conveys two main pieces of information. First, it is clear that the absolute magnitude of frequency-domain GC values are stronger in the vStr→OFC direction, than vice versa, in agreement with the data from panel A. Second, it shows the main spectral peaks for G-causal interactions between OFC and vStr. There are noticeable peaks for both directions at beta (~25Hz) and gamma50 (~50Hz) frequencies, and a peak in the gamma80 range (~85 Hz) for the vStr→OFC direction. Overwhelmingly though, the dominant frequency for both structures is at gamma50, again with vStr→OFC being larger in absolute magnitude. Taken together, the data in Figure 4-10

strongly indicates that on the delay discounting task, it is the ventral striatum that is chiefly driving the orbitofrontal cortex, specifically in the gamma frequency range.

The single unit and ensemble decoding data from Chapter 3 demonstrated that both OFC and vStr showed a strong and sustained response to reward on the DD task. To compare the dynamics between OFC and vStr with the within-structure analyses in Chapter 3, PETHs of time-domain GC value were constructed, aligned to the same task events. Figure 4-11 A shows the causal influences between OFC and vStr around the time of reward receipt. GC values are consistently higher in the vStr→OFC direction than vice versa, consistent with the session-level data in Figure 4-10. Apart from this difference in magnitude, both directions of causal influence increase sharply after the time of feeder fire, starting about 2 seconds after the feeder trigger event, at the time when the food pellets arrive. This result is interesting in that OFC and vStr appear to both be exerting influence. Although we know that single cells and ensembles in both structures are responsive to reward receipt, we might not expect *a priori* that both structures would show an increase in their GC values. Parallel increases in GC values aligned to stimulus events are have been reported. See [219] (Figure 3) for an example from prefrontal and posterior parietal cortex. The data from Figure 4-11 A shows both OFC and vStr exhibit an increase in directed interactions after the salient reward delivery event.

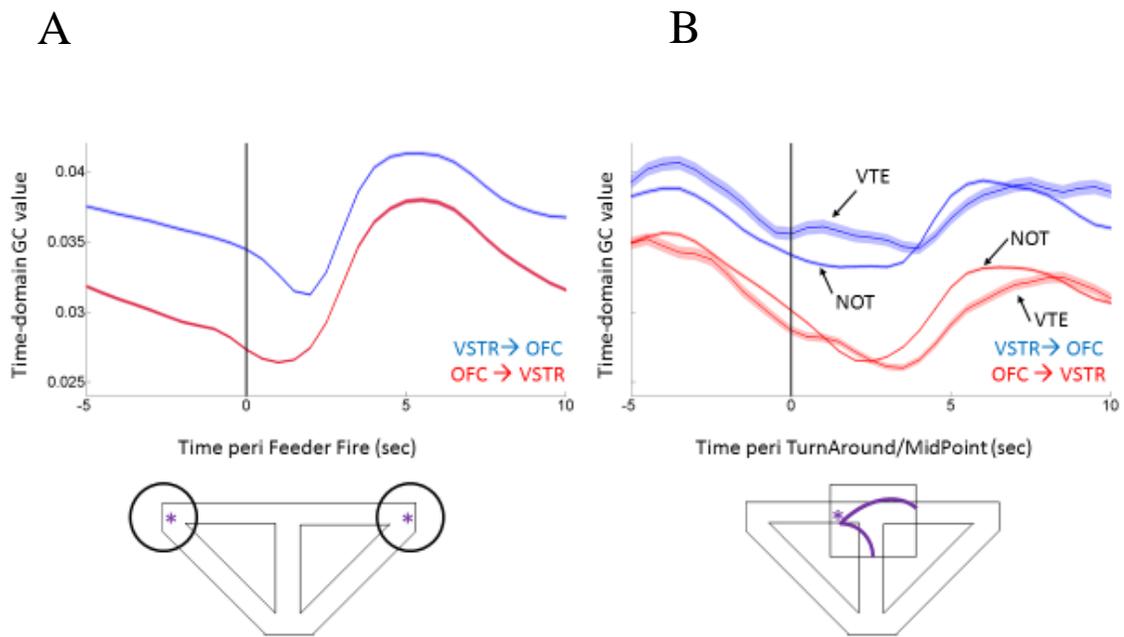


Figure 4-11. **Granger causality at the choice point and during reward receipt.** Panel (A) shows time-domain GC values around the time of Feeder Fire. GC values increase markedly at ~2 sec after Feeder Fire, the approximate time of reward receipt. Panel (B) shows time-domain GC values aligned to the TurnAround Point (VTE laps) or the MidPoint (non-VTE laps). Note the increased values for VTE laps in the vStr→OFC direction.

Decoding analyses at the choice point showed that vStr and OFC ensembles represented the feeder site during VTE, with the vStr preceding OFC, and occurring before the moment of choice in the vStr (Figure 3-9). How is information transmitted between OFC and vStr during the decision-making process? Figure 4-11 B shows time-domain GC values for OFC and vStr, aligned to the point of TurnAround or to the MidPoint, for VTE and non-VTE laps, respectively. As before, the absolute magnitude of GC values are greater in the vStr→OFC direction than vice versa. And as before, GC values in both directions have the same general shape. However, GC values from vStr→OFC are increased during VTE laps as compared to non-VTE laps, and this effect is not present in the OFC→vStr direction, or perhaps even reversed at around the five second mark (NOT greater than VTE). These data indicate that information transmission from vStr→OFC is greater during VTE laps, in line with the decoding results in Chapter 3.

Interestingly, this difference between VTE laps and NOT is most marked *after* the point of TurnAround, which also corresponds to the time when reward-site decoding in OFC increased. The result in Figure 4-11 B raises the possibility that the vStr is providing inputs that contribute the “bump” in feeder decoding in OFC. However, that question cannot be answered with this analysis. Granger causality measures the degree to which signal A is predictive of signal B, but it does not specify what kind of information is being transmitted, at least in terms of neural representations. The data here is consistent with a special role for the vStr during deliberative behavior, but the exact nature of what information is being transmitted requires further investigation.

A second and complimentary analysis was applied to the same paired LFP data to estimate the directionality and lag between OFC and vStr LFPs at different frequencies. This technique, which measures the cross-correlation between the two LFP signals, filtered at a desired frequency range, provides a more mathematically intuitive measure of directionality [171]. Briefly, the two LFP signals are filtered, their cross-correlation is measured, and then the maximum value of the cross-correlation curve is taken as an estimate of the lag between the two signals. If the peak of the curve falls to one side of the zero line, it indicates that signal A precedes signal B. If the peak falls on the other side of the line, it indicates that signal B precedes signal A. Taking the distribution of these maximum cross-correlation values across multiple observations provides an estimate of the directionality and the lag between the two signals.

This maximum cross-correlation measure was calculated for each session ( $n = 140$  sessions) and the distribution of maximum cross-correlation values for each of four different frequency ranges is plotted in Figure 4-12. These data are broadly consistent with the Granger causality analyses just described. For most frequencies, with the exception of theta, the vStr leads the OFC, suggesting that the whole session level, the direction of information transmission is from vStr to OFC, in line with Figure 4-10 A. The data is highly significant for the gamma frequency range (sign test,  $p \ll 0.05$ ), also in agreement with the Granger causality analysis (Figure 4-10 B).

The fact that the OFC appears to influence the vStr at theta frequency is interesting. Significant theta power has been reported in both striatum [119, 147, 220] and OFC [221] (albeit much lower in amplitude in both cases as compared to the hippocampus). However, theta power in the vStr and OFC was not apparent in the power analyses that we used, which were averaged over the whole session. Theta power might be more evident at high running speeds, which has been shown in mPFC [196], or at specific task epochs, like odor sampling [221]. The relationship between OFC and vStr in the theta range merits further investigation. Overall, the cross-correlation analysis provides confirmatory evidence that, broadly speaking, the vStr is transmitting information to the OFC on the delay-discounting task.

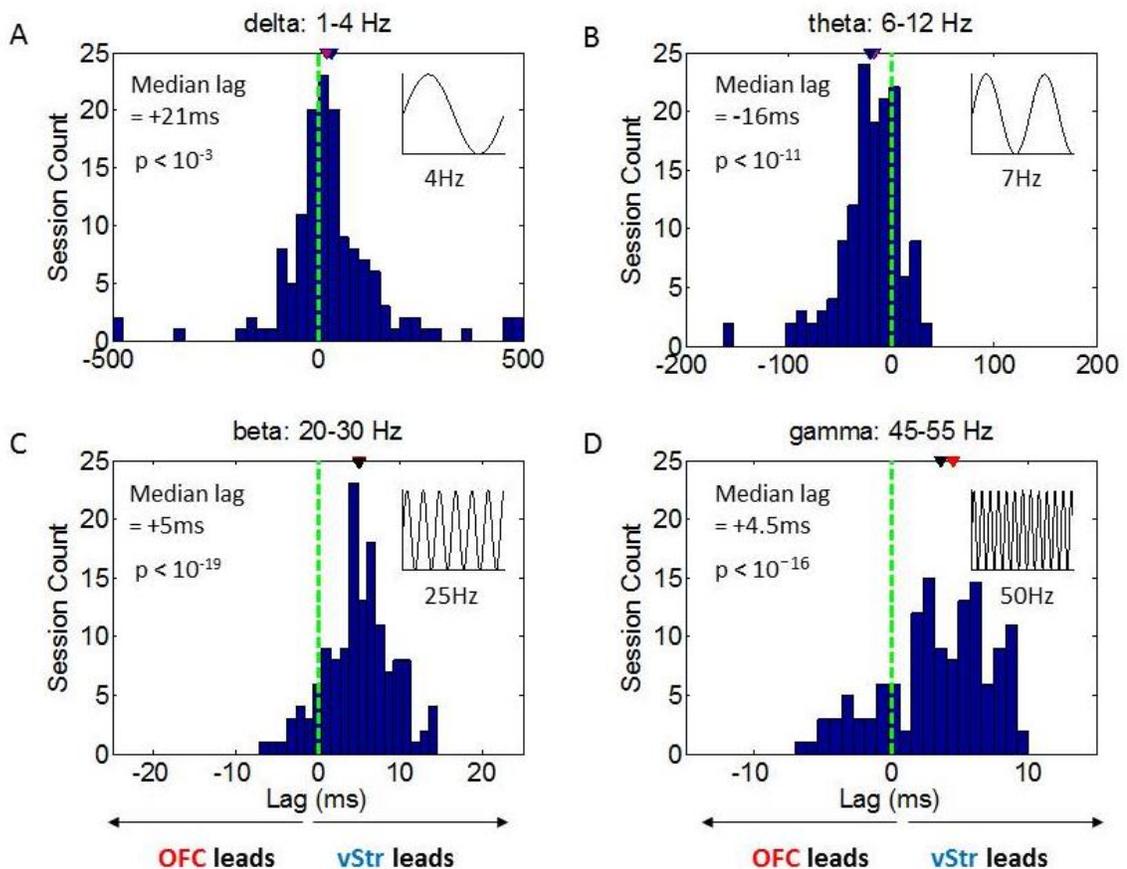


Figure 4-12. **Maximum cross-correlation distributions.** Each plot shows the distribution of maximum cross-correlation values across sessions ( $n = 140$  sessions), as a function of lag.

Frequency ranges are **(A)** delta (1-4 Hz); **(B)** theta (6-12 Hz); **(C)** beta (20-30 Hz); **(D)** gamma (45-55 Hz). Black lines are at zero and red lines indicate the median of the distribution. All distributions had a median value significantly different from zero (two-sided Wilcoxon signed rank test,  $p < 0.05$ ). Red lines to the right of zero (positive values) indicate that the vStr leads OFC, and red lines to the left of zero (negative values) indicate the OFC leads vStr. The x-axis differs between plots to capture the minimum and maximum values, while aiding visualization. Median lags were 21ms, -16ms, 5ms, and 4.5ms for delta, theta, beta, and gamma, respectively. Thus, the vStr led OFC in all frequency ranges except for theta. The distribution for gamma80 (70-85 Hz) was also highly significant, with a lag of 2.51ms.

The cross-correlation measure complements the GC analysis in that it provides an explicit estimate of the lag between structures, which is not provided by G-causality. The median lag for the cross-correlation measure was 6.5ms. This lag would appear to be short, given the polysynaptic relay between the striatum and OFC. However, several studies have described coherent activity between non-adjacent areas at short latencies, specifically in the gamma band, as seen in my data:  $V1 \rightarrow V4$ , only a few milliseconds [222];  $FEF \rightarrow V1$ , ~10ms [223]; inferior frontal junction (IFJ) to parahippocampal place area (PPA), ~20ms [224]. Our data are also consistent with a recent report that used a different measure, called “phase-slope analysis” [225], to estimate the directionality and lag between ventral striatum and medial prefrontal cortex [226](the lag there was even shorter). Filtering between 40-70 Hz, the lag was 6.5ms, which may fall within the range of conduction delays between vStr and OFC. Zero-lag or near zero-lag phase relations between structures have been recorded in visual cortical areas [227], and these can be the result of bidirectional interactions [228], which could apply to the OFC and vStr (as parts of a basal ganglia cortico-striatal loop).

Beyond the time lag, another question to consider is how coherent activity can be maintained between structures separated by several anatomically different nuclei. In other words, can oscillatory activity at a specific frequency travel through the basal-ganglia-cortical loop? Our coherence and directed connectivity measures suggests that it can. A recent study supports our findings. In paired recordings made specifically in the striatum (both dorsal and ventral aspects)

and the PFC on a category learning task, Granger causality analyses revealed significant and stronger causal connectivity from the striatum to PFC than from the PFC to striatum in the beta band [229] (Figure 6). This result provides confirmatory evidence that frequency-specific GC influences can be transmitted between polysynaptically connected brain regions in the BG-cortical loop.

## Discussion

Analyses were performed on simultaneously recorded local field potential signals from the orbitofrontal cortex and ventral striatum in rats performing the spatial adjusting delay-discounting task. An analysis of the power spectra in each structure revealed fundamental frequencies at gamma50 for both OFC and vStr, and additionally at gamma80 for the vStr. There was noticeable power in the beta frequency range (~25 Hz) in OFC and vStr, as revealed by self-coherence plots, with a significant correlation between beta and gamma50 for both structures. The self-coherence plots for vStr also revealed an anti-correlation between gamma50 and gamma80, consistent with a previous report [119], and illustrated by alternating bouts at these frequencies in the spectrogram shown in Figure 4-4. Altogether, these data are consistent with other LFP recordings from the OFC [129] and vStr [128]. In particular, these data affirm the presence of two distinct frequency bands within the gamma range in vStr [119].

Around the time of reward receipt, the OFC and the vStr had very similar and characteristic patterns of gamma power modulation, but this modulation differed between gamma50 and gamma80 sub-bands. Gamma50 power increased sharply after the feeder trigger event, while gamma80 power sharply decreased. A previous study found evidence that gamma50 power in the vStr discriminated between reward sites and between different types of reward [149]. On the DD task, vStr gamma50 power was elevated before the large reward (as compared to the small reward), but this was not evident at gamma80. Overall, these patterns of gamma50 and gamma80 modulation around the time of reward receipt are consistent with previous reports [119, 149].

The main theme connecting these data is that gamma50 and gamma80 power showed differential activity dynamics. The same was true for coherence in the gamma50 versus the gamma80 range. Gamma power and coherence in these two frequency bands alternated at the time of reward receipt. At the choice point, gamma50 power in vStr and gamma50 coherence was elevated after TurnAround events. How does this relate to our understanding of gamma oscillatory activity during reward-guided behavior?

In the paper by van der Meer & Redish [119], gamma50 power was flat during the navigation sequence (through the central stem and goal arms) and increased after arrival at the feeder site. Gamma80, on the other hand, “ramped up” as the rats’ ran down the track and peaked just before arrival at the feeder site. Gamma50 power at the reward site increased over laps [119]. Although we did not measure gamma power along the track, we did find that gamma50 peaked after reward receipt, while gamma80 peaked earlier, at the time of Feeder Fire (corresponding roughly to arrival at the feeder site). These data provide suggestive evidence that gamma80 may be more involved in the anticipation of reward, while gamma50 may code for reward receipt itself. The vStr has been proposed to send “state value” signals to the VTA [77], indicating the value of the present state the animal is in. State value would be expected to increase with proximity to reward. If the vStr does provide these kinds of value signals, this could be mediated by gamma80 ramping activity. Gamma80 power in the vStr increases dramatically after dopamine agonist administration [148], and tonic dopamine in the vStr also “ramps up” as rats run along a track, approaching reward [230]. Thus, gamma80 power may signal reward-expectancy in a dopamine-dependent fashion. It would be interesting to see if phasic dopamine bursts increase gamma80 power.

In contrast to this view, our data did show higher gamma50 activity in the vStr before reward on delay side laps (i.e. during reward anticipation). Thus gamma50 may play a role in reward anticipation as well, as least during sustained delays. However, activity during the adjusting delay was universally high in the decoding and population firing rate PETHs as well. Waiting at

the feeder site may involve several cognitive/behavioral components besides reward anticipation, including sustained attention, inhibiting alternative motor actions (i.e. leaving), and time estimation. Therefore, we cannot relate activity during the delay solely to reward anticipation or prediction (and this is true of other studies).

The data here and in the literature show different patterns of activity at gamma50 and gamma80 frequencies, especially in the vStr. This provides tantalizing evidence that LFP activity in the vStr (and OFC) might play a role in influencing behavioral output. Gamma coherence between OFC and vStr followed the same pattern of activity seen in the gamma power data. If gamma50 and gamma80 power have distinct behavioral functions, this could be mediated by interactions between the OFC and vStr. Causal manipulations of activity in these structures, or lesions of either the OFC or vStr could shed light on this question.

The coherence analyses indicated that there were specific frequencies and specific times during the task where functional connectivity between OFC and vStr was high, indicating possible communication between these structures. Coherence, however, does not provide information about which structure is driving the other. Granger causality (GC) analysis was applied in order to determine which structure exerted a greater influence on the other during the delay-discounting task. As noted at the beginning of this chapter, the results from Chapter 3 demonstrated that the vStr preceded the OFC in representing information about future reward and information about the rat's impending decision. We asked whether this difference in timing might be reflected in the directed connectivity between OFC and vStr. We found that this indeed was the case.

At the whole session level, a clear pattern emerged. For a majority of sessions, the vStr had a stronger influence on the OFC than vice versa. In the frequency domain, the absolute magnitude of GC values was also much larger for vStr→OFC than vice versa, and this effect was concentrated in the gamma50 frequency range. These data indicate that both the OFC and the vStr have significant causal influences upon one another, particularly in the gamma50 frequency range, but that the stronger driver during the delay discounting task is the ventral striatum. These data were

paralleled in analyses of specific task epochs, where vStr→OFC GC values were stronger than OFC→vStr GC values at the choice point (Figure 4-11 B) and in response to reward receipt (Figure 4-11 A). Importantly, at the choice point, GC values from vStr to OFC were greater during VTE laps than during non-VTE laps, and this effect was absent (or even reversed) in the OFC to vStr direction (Figure 4-11 B). This suggests that the vStr has greater influence over the OFC during deliberative decision-making. How does the timing of this effect line up with the results in Chapter 3?

In the decoding analyses, covert reward decoding occurred *before* the moment of choice in vStr. Here, Granger causality from the vStr to OFC was stronger *after* the moment of choice. These results may seem at first to be at odds with one another. If the vStr is transmitting reward-related information to the OFC during VTE, then we would expect this to occur at the same time that the vStr starts to show increased feeder site decoding. However, given the fact that there is some transmission delay, and that the temporal precision of these analyses is not at the millisecond timescale, a slight delay is not unreasonable. Alternatively, the vStr might transmit information that is not identical to that captured in the decoding analysis. This remains to be determined.

The results from the maximum cross-correlation analysis (Figure 4-12) were broadly consistent with the Granger causality analysis, showing significant and stronger information transmission from vStr to OFC in three of the four frequency bands tested. The distribution of values at gamma frequency favored vStr leading OFC, and was highly significant, in line with the spectral values in the gamma range shown in Figure 4-10. Altogether, these, data support the interpretation that the vStr is a strong driver of OFC activity during an economic decision-making task. Moreover, this effect was greater during deliberative VTE events. These results run counter to the standard notion that the cortex sends top-down signals to the basal ganglia in order to affect executive decision-making [231]. While that may be the case in many circumstances, we have found that the vStr can also be the primary driver of activity in the cortex, particularly during deliberation.

## Chapter 5: Discussion

### Summary of results

We performed simultaneous neural recordings from the orbitofrontal cortex (OFC) and the ventral striatum (vStr) on a novel decision-making task, the spatial adjusting delay-discounting task. The central findings are as follows: Behaviorally, rats showed economic behavior on the DD task, along with frequent VTE events—a signature of deliberation—that matched the transition from flexible to automatic response strategies. At the single unit and ensemble levels, OFC and vStr showed overlapping activity, but distinct differences in the timing of decoded reward-related and choice-predictive activity, with vStr leading OFC. At the level of local field potentials, OFC and vStr showed prominent gamma oscillations that were modulated by task parameters, coherent activity in the gamma range, and a strong bias in directed connectivity, with the vStr exerting a stronger influence over the OFC at the session level, and, notably, increased causal influence during VTE. Below I discuss the implications of these results, and their relation to the literature on VTE and the roles of the OFC and vStr in decision-making.

### VTE on the delay-discounting task

Rats on the spatial delay-discounting task consistently titrated the adjusting delay to a preferred level, showing that they made a tradeoff between reward magnitude and delay to reward (Figure 3-2). In an experiment that systematically changed the magnitude of reward at the two feeder sites, rats on the DD task linearly increased their willingness to wait as a function of the reward ratio, consistent with hyperbolic discounting [112] (Figure 2). These data demonstrate that rats on the DD task engage in temporal discounting, a well-studied form of economic behavior. Rats also showed frequent VTE behavior on the DD task (Figure 3-3, Figure 2-6 for examples). VTE has been linked to deliberation and future planning [69, 113], making the spatial delay-

discounting task a useful tool for studying the interaction between deliberative processes in the rodent and economic decision-making behavior.

A central premise of this thesis is that VTE is a behavioral marker of deliberation in the rodent. So it is worth examining this position. Does VTE reflect deliberation, or is it a manifestation of some other process? The simplest criticism of VTE is that it is merely an epiphenomenon of some other process, unrelated to decision-making. Guthrie criticized Tolman's proposed mechanism of VTE as leaving the rat "buried in thought" at the choice point [232]. This simple criticism is contravened by a large body of evidence showing that VTE occurs at times and places that are in fact the most relevant for decision-making. Converging data from pharmacological and neural recording studies strongly support the idea the VTE is behaviorally relevant and involves neural activity that is involved in decision-making [48, 114, 120, 180, 233-238].

#### Alternative accounts of VTE behavior

However, there are non-trivial alternatives to the deliberation hypothesis that could hypothetically explain VTE behavior. As discussed in Papale et al. [112], these alternatives include conditioned orienting, perceptual or value discrimination, or exploration. Each of these alternatives are taken up in turn.

VTE is indeed prominent in perceptual discrimination tasks [116, 236, 239] where explicit sensory cues indicate correct versus incorrect choice options. Conditioned associations to these affectively charged cues<sup>12</sup> could elicit Pavlovian approach and avoidance behavior, which might manifest itself as hesitation at the choice point. Thus, VTE behavior could reflect looking back and forth between different cues to gather sensory information, or it could reflect conditioned orienting (i.e. Pavlovian approach).

---

<sup>12</sup> Some early studies of VTE paired responses with shock [239], or required the rat to jump over a significant gap to receive reward on the adjacent platform. Jumping toward the incorrect platform caused the rat to fall into a net [116].

Once the rats exited the choice point, tones provided sensory feedback about the chosen option. Anticipation of these cues could in theory drive a Pavlovian response (i.e. conditioned orienting). If VTE resulted from conditioned orienting, then VTE should be more frequent at shorter delays, because the associative strength of conditioned stimuli is greatest when the CS-US interval is short [240]. This was not supported by the data. In fact, the opposite was true. VTE was most frequent at longer delays (see Papale et al. [112], Figure 7). Thus, conditioned orienting to the reward-paired cues on the DD task does not account for the timing of VTE events within the session.

VTE is also not well accounted for as a sensory discrimination process. On the DD task, proximal and distal cues are fixed across sessions, while the number of pellets obtained at each feeder is counterbalanced across days. Therefore, the sensory properties of the maze do not provide consistent landmarks for action selection. In a walled version of the spatial delay-discounting maze [241], rats engaged in VTE behavior at the choice point, even though they could not see the reward sites—presumably, the most salient visual cues on the maze. Likewise, rats increase their frequency of VTE behavior after a reversal in reward contingency, even though nothing about the maze or the distal cues in the room have changed [83, 120]. Neurally, hippocampal representations during VTE often “sweep” down the maze arm opposite to the arm that the rat is facing [114], suggesting that VTE involves an internal cognitive event that can (at least in some cases) be uncoupled from sensory information. Thus, a purely sensory explanation does not account for the range of conditions under which VTE occurs.

An alternative theory is that VTE involves a sensory-like discrimination problem, but instead estimating value. This notion draws a parallel to sensory discrimination studies in which animals are known to accumulate sensory information—up to a threshold level—in order to improve their performance [242]. If VTE were an instantiation of an integrate-to-threshold process, then VTE should be highest when the expected values of the two choice options are closest together. Value discrimination was not supported. VTE was lowest when the values of the two reward

options were the same (i.e. during the exploitation phase). However, VTE could play a role in value discrimination before decision-making has become habitual, and indeed, that may be the case early on in the DD task. However, neural evidence for such a process has proved elusive.

Lastly, VTE could reflect exploration of the environment. Rats show pausing and rearing behavior when investigating a new environment. However, simple exploration is unlikely to account for VTE, because rats were already very familiar with the environment before the start of the recording sequence. VTE, however, could reflect the “exploration of information.” This is similar to Tolman’s account, in which VTE is the internal (vicarious) exploration of alternatives (trial and error) when rats face uncertainty in the environment.

In line with Tolman’s conception, Johnson et al. [72] have put forward a normative account of rodent exploratory behavior that explains VTE in terms of “information foraging.” Here, VTE is quintessentially about gathering information, by means of sampling internal memories. This theory predicts that “directed exploration” (and consequently, VTE) should occur when rats are familiar with the environment and understand the overall task structure, but still have some uncertainty with regard to the value or likelihood of reward among the different choice options. If VTE involves directed information foraging on the DD task, then VTE should be low at the very beginning of the session. This is because the rat would start out with no preconceived idea about which side has the small versus the large reward, or the length of the adjusting delay. In other words, rats should start out with a uniform prior<sup>13</sup> belief state. At this time, rats likely use a random, trial and error sampling strategy [72]. VTE should increase when the animal has experience with the specific task parameters for that day’s session, but still has some uncertainty about the reward options, and therefore can *gain information* (because the animal now has a prediction that can be

---

<sup>13</sup> The use of the word “prior” here is different from the spatial prior referred to in the decoding analyses of Chapter 3. Both terms are Bayesian “priors” and are, quantitatively, a probability between 0 and 1. The spatial prior refers the likelihood of the rat occupying any given position on the maze. A “prior belief” about reward, in the model used by Johnson et al. [72], is the likelihood that a given feeder port is the correct (i.e. rewarded) option.

tested). Once the animal has settled on a certain belief state, VTE should fall off. Thus, VTE has three phases, just as Tolman believed [243].

The three phases of “directed information foraging” (see Figure 4 of [72]) match up with the behavioral data reported in this manuscript. During the recording sequence (from which the behavioral data here was taken), rats on the DD task were familiar with the overall task structure (a necessary requirement for directed foraging). However, at the start of each session they did not know which side offered the large reward nor did they know the initial delay (introducing uncertainty). Rats began the task with alternation behavior (investigation phase), which may be considered a default strategy [244, 245], in line with the idea that rats start off with random sampling of the choice options [72]. Rats then engaged in titration of the adjusting delay (titration phase), during which time the waiting period for the larger-later reward was changing. This meant that the rats likely had uncertainty about the adjusting delay during this period, and therefore could gain information by sampling the opposite side. Most titration phases on the DD task were “downward titration” sequences, because the initial delay (average of 15 sec) was more often above rats’ indifference point (~6 sec) than below it (Figure 3-2). Thus, rats favored the non-delay side early on to “drive down” the adjusting delay. But they did often sample the delayed side before reaching their indifference point, perhaps to gain information about the length of the adjusting delay. Finally, rats settled on an alternation strategy during the exploitation phase, during which time they experienced a constant delay on both sides of the maze. This would correspond to a state of high certainty and less need for directed exploration.

The timecourse of VTE during the session matched these behavioral dynamics. VTE was low initially<sup>14</sup> (investigation phase). VTE rose during early laps, reaching a peak at about lap 20 (titration phase). The frequency of VTE then gradually decreased and settled to a baseline level (exploitation phase). These three phases can be interpreted in information theoretic terms [72], as

---

<sup>14</sup> Apart from lap 1.

discussed above. Thus, the VTE analyses here are consistent with the evolution of VTE behavior predicted by the information foraging account. However, in the larger dataset in Papale et al. [112], VTE was highest on the very earliest laps (Figure 6b of that paper). VTE was high at the very outset across all three behavioral phase classifications (Figure 8 of that paper). The reason for this difference is unclear. We calculated VTE frequency as a function of lap using the zIdPhi measure employed by [112], and this plot was essentially the same as Figure 3-4 C, which used curvature to define VTE events. Therefore, this discrepancy is not due to the algorithm used to classify VTE. Regardless, the general trend that VTE is higher early in the session than late is in agreement with the idea that information sampling would be necessary early on, and not so late in the session. That being said, VTE did occur even during the latter part of the session (it did not fall to zero), and this fact is not explained by the information foraging model<sup>15</sup>.

Both the deliberation (i.e. Tolmanian) and information foraging theories entail mental simulation of some aspect of alternative actions (be it reward magnitude, time to reward, or some other aspect of the task), and both predict VTE after a reversal or change in contingency. As mentioned above, both predict three phases of VTE behavior. The primary difference between these two accounts is that information foraging is based on the idea that the animal is trying to maximize information. Under the information foraging account, the animal selects actions which will yield the greatest *increase* in information about the different reward options (when there is uncertainty). It relies on *past experience* to perform an internal search for the most informative option. Deliberation, on the other hand, occurs in situations where gaining information is not the objective, and where the decision is entirely novel. A person or an animal may well deliberate over a “one-shot” decision which has no prospect of information gain. Many important decisions—like where to go to college—are one-shot decisions. Thus, a task that involved one-shot decisions, or novel

---

<sup>15</sup> In reinforcement learning models, choice is partly stochastic, even during asymptotic performance. The agent is essentially checking the other options to make sure nothing has changed. This could explain the persistence of VTE throughout the session.

decisions, could in theory distinguish between information foraging and deliberative accounts of VTE behavior.

#### VTE reflects purposive behavior

There are many lines of evidence that support the claim that VTE reflects purposive behavior, and in particular, that VTE involves a deliberation process. Importantly, across studies, the distribution of VTE events is not random. In my data, the frequency of VTE events was significantly higher during the first half of the session, when the adjusting delay was changing, then during the second half of the session, in which rats alternated sides, using a fixed exploitation strategy (Figure 3-4). The first half of the session was characterized by titration behavior. Thus, the rats may have been intentionally changing the adjusting delay in order to reach a long-term objective. Additionally, titration of the adjusting delay involves repeatedly going to the same side, which is opposite to the natural tendency that rats have to switch or alternate sides [244, 245]. Thus, returning to the same side on adjustment laps likely means that rats were inhibiting a prepotent response tendency, an example of executive control. These factors argue for the interpretation that during the titration phase of the task, rats engaged in goal-directed behavior.

The relationship between VTE and choice on the DD task also argues for a deliberative interpretation. On non-VTE laps, rats almost exclusively chose the opposite side goal arm (alternation). However, on laps in which VTE occurred, rats' chose the opposite-side and the same-side goal arms with equal likelihood (Figure 3-5). Thus, during non-VTE laps, knowing where the rat was coming from provided a high degree of information about which side the rat was going to choose; during VTE laps, knowing where the rat was coming from provided no information about which side the rat would choose. Rats were essentially undecided on VTE laps, suggesting that they made their choice "on the fly" at the choice point, consistent with a deliberative process. This finding was paralleled in the neural decoding data. The difference in decoding strength between the chosen and non-chosen feeder sites rose above chance earlier on non-VTE laps than on VTE

laps (Figure 3-14, a vs. b). This is consistent with the result above; namely, that animals seemed to have “made up their minds” earlier on non-VTE laps. On VTE laps, neural activity differentiated the animals’ choice later, as if the rats were slower in committing to their decision, perhaps due to the processing time required for deliberation.

An important line of evidence for the role of VTE comes from work on the dorsal striatum, which is closely tied to the formation of habits—at the other end of the spectrum from goal-directed or flexible behavior. When rats are well trained on a T-maze task, spiking activity in the dorsolateral striatum (DLS) becomes concentrated at the beginning and end of maze runs [79, 81, 83], a phenomenon known as “chunking”, or, synonymously, “task-bracketing”. Chunking activity in the DLS is thought to facilitate automatic and efficient use of well-learned stimulus-response associations (S-R) [246]. In a recent experiment that used reward devaluation to measure habitual responding [82], task-bracketing activity in the DLS was shown to be inversely correlated with the likelihood of VTE on a lap by lap basis. Additionally, VTE was almost entirely driven by laps in which animals looked toward the instructed, but devalued reward arm and then went in opposite direction. These VTE events faded as a new habit formed. Thus, VTE occurred specifically on laps in which there was a conflict (making evaluation advantageous), VTE frequency decreased as performance became habitual, and VTE was *inversely* related to task-bracketing activity in the DLS, a neural signature of habitual responding.

In a different experiment, this same basic pattern was replicated at the single session level, with VTE increasing and task-bracketing decreasing after a contingency switch, and VTE falling off and task-bracketing re-emerging as the new contingency was learned and behavior automated [83]. Importantly, significant task-bracketing was not found in the hippocampus [83], a region associated with flexible navigation [73, 74], future planning [118, 176], and VTE [114]. Taken together, data from the DD task [112, 144, 241], previous (discussed in Chapter 1: Introduction) and current work from our lab [48, 114, 120, 180], and work from other laboratories [82] argue that VTE reflects purposeful, deliberative behavior.

### Reward responsivity on the DD task

At the single unit level, activity dynamics around the time of reward receipt were very similar between OFC and vStr. Both structures had a large fraction of reward-responsive neurons (70-75%), and had similar numbers of cells that increased or decreased their firing rate in response to reward (Figure 3-6). The fraction of reward-responsive units was on par with other studies [48, 120], but toward the high end. One factor to consider is that we included cells that decreased their firing rate in response to reward, thus increasing our pool of reward-responsive units. This subset of cells is often omitted for analysis [37]. The ensemble decoding (Figure 3-11) and population firing rate dynamics (Figure 3-12) around the time of reward receipt showed phasic peaks after the reward cue and at the time of reward receipt, consistent with the literature, which shows that the OFC and vStr encode both expected and received reward[52]. Activity was also maintained during the waiting period (on delay side laps) (Figure 3-12), as reported by others [183, 247]. Thus, activity in OFC and vStr was very similar, although further analyses could reveal subtle differences.

An interesting aspect of this data is the persistent activity in OFC and vStr during the adjusting delay. Persistent activity is not likely to reflect a mere timing mechanism (a clock), as it did not ramp up, but rather stayed level. A more interesting possibility is that the persistent activity is involved in working memory. As has been shown for delayed-response tasks in primates [248], persistent firing could reflect the operation of a working memory buffer to bridge a temporal gap. Instead of encoding a memory trace for a spatial position, persistent activity here could be necessary or involved in maintaining sustained attention to the feeder site—staying vigilant to consume the reward as soon as it appears. Dopamine receptor activation in the NAc core is necessary for “flexible approach behavior”, which in this case, would involve returning to the feeder port if the rat is distracted or moves about during the delay [109]. Dopamine is also critically involved in motivation and the invigoration of behavior [249], and for facilitating the ability to respond to unpredictable stimuli [250]. Although the rats could know the length of the adjusting delay on a

given lap based on past experience (and knowledge of the task), they likely have some uncertainty at the very least due to the estimation error inherent in judging intervals. Dopamine-evoked accumbens activity could therefore be involved in sustaining attention during the delay. The role of the OFC in this context is not as well studied, but the OFC, like most of the prefrontal cortex, does have significant dopamine receptor expression.

Another possibility, not exclusive with the first, is that persistent activity is needed to remain at the feeder in the face of a competing urge to leave the feeder. Selecting any action involves an opportunity cost. Rats on the DD task always have the option of leaving the delayed side prematurely, forfeiting the larger reward in favor of running a lap to the immediately rewarded side. Serotonergic transmission in the OFC has been linked to impulsivity and the ability to persist through a delay [251, 252]. However, a recent report found that 8-OH-DPAT—a 5-HT1A agonist—injected into the OFC, led to *increased* impulsive choice on an adjusting delay task [253]. The effects of serotonin are likely to be sub-type specific [254]. Real time neurochemical monitoring techniques would prove useful in investigating this question. In practice, rats on the DD almost never skipped at either of the feeders, precluding a detailed analysis. They did skip on rare occasions, usually on early laps when the initial delay was high. This question could be addressed with a task that elicits frequent skip trials that occur over a range of delays.

### Value integration

Economic theories posit that humans integrate all relevant decision variables (size, delay, effort, cost, probability, risk, etc.) into a single subjective value before comparing options [255]. In neuroeconomic models of choice, for tasks that involve a tradeoff between reward magnitude and delay, these factors should be integrated into a singular temporally discounted value signal [9]

Behaviorally, rats on the DD task did show temporal discounting (Figure 3-2; [112]), as economic theory dictates, and as has been shown in a plethora of animal studies of delay

discounting [256]<sup>16</sup>. Neurons in OFC and vStr showed very strong responses to reward and reward-predicting cues on the DD task. Both the OFC and vStr have been strongly implicated in delay discounting. How did these neurons change their firing rate response as a function of reward magnitude and delay?

We found relatively modest levels of magnitude and delay coding in OFC and vStr (~20% of cells). However, these numbers are in the same ballpark as other neurophysiological studies (discussed in Chapter 3). However, we did not observe any correlation between magnitude and delay coding at the single cell level (Figure 3-10), arguing against temporally discounting value in a common-currency scheme.

How do these results accord with the literature? Some studies in the primate have found integration of economic variables in single cells [17, 183], while others have not [28, 258]. In the case of temporal discounting, the absolute number of large-preferring and short-preferring neurons in these studies is not large (about 20% each, as discussed in Chapter 3), and the overlap represents a very small fraction of the population [259]. Thus, the empirical data from primate neurophysiological recording studies has shown mixed results, at best, for true value integration in either the OFC or vStr. Rodent recording studies have consistently failed to find value integration at the single cell level [260], and additionally, have found sensorimotor modulation of value signals in OFC [37, 38], inconsistent with a candidate abstract value signal. Apart from a few influential examples [17], the bulk of the neurophysiological data does not support integrated value signaling

---

<sup>16</sup> Rats showed economic behavior in the sense that they discounted delayed rewards, and in a consistent manner. However, economic theory postulates that the particular shape of the discounting curve should be exponential. In an experiment that systematically manipulated the reward ratios on the DD task, rats' behavior was consistent with hyperbolic discounting [112]. Exponential discounting is rational in the sense that it involves a constant discount rate over time. For each increment of time, the value of the delayed reward is reduced by a fixed percentage. For example, if the value of your car depreciated by 20% every year, this would be an exponential reduction in value with time. Anything other than exponential discounting, such as hyperbolic discounting, results in preference reversals. An agent will prefer a larger-later reward when it is considered at a long time interval, but will change its mind as the time to reward nears. However, the vast majority experiments in human and animal subjects favor hyperbolic discounting as a better fit to the data [257].

in a common-currency format. Thus, how delay-related and magnitude-related information are integrated—if indeed this is the case—remains an open question<sup>17</sup>.

Whether value integration needs to happen in a single brain area, or if it needs to occur at all, is a hotly debated topic [35, 260, 261]. Alternative accounts, such as the distributed consensus model [262], and the multiple systems model of decision-making [64, 69], do not require abstract value signals in the strict sense. In the latter model, emphasis is placed on how different contexts, task-demands, and the extent of experience influence which action-selection systems are employed.

### Multiple decision-making systems

The multiple-systems model of decision-making predicts that value signaling should depend on the action-selection system being used [64]. For the DD task, it predicts that reward-related signaling should differ based on whether the deliberative system or the habitual system is exerting more control. Specifically, it predicts that reward-related signaling should be present at the time of choice when the animal is engaged in deliberation, because deliberation involves a search process that includes estimating the value of potential actions [76]. During habitual behavior, reward evaluation should be relatively weak at the choice point, since the animal is using a stimulus-response strategy based on previously learned cached values (instead of prospective valuation signals).

Indeed, this is what we found. Decoded reward representations were higher in both the vStr and the OFC around the time of choice (TurnAround) on VTE as compared to non-VTE laps (Figure 3-13). Signals consistent with the evaluation of future reward were present at the choice point during VTE behavior, as would be expected if VTE reflects a deliberative process with a reward evaluation component [58]. Decoded reward activity in the vStr started increasing *before*

---

<sup>17</sup> This is not to say that magnitude and delay-related information is not involved in the decision-making process. Clearly, they are necessary pieces of information in performing tasks where these variables are involved.

the time of TurnAround, after which time the rat engages in ballistic movement toward the goal arm and has committed to his decision. This data reinforces the notion that the vStr can generate expectancies of future rewards in the absence of external cues, before the time of choice [48](Figure 10 for timing), linking the vStr to the flexible decision-making system. Overall, the data in this thesis is consistent with the existence of multiple decision-making systems that recruit anatomically distinct brain regions, with the OFC and vStr being part of the deliberative system [69].

### Implications of the decoding results

Expectancy related signals were present during VTE on the DD task in the vStr (before the moment of choice) and the OFC (after the moment of choice). The term “expectancy” is used here to mean a representation of the outcome before it has occurred. Once a matter of debate, it is now generally accepted that animals generate expectancies to inform their decisions [59, 263]. Expectancies consist of (1) identifying possible actions, and (2) evaluating the outcomes associated with those actions. Expectancies are a critical component of “model-based reinforcement learning” (alternatively, “model-based decision-making”). Expectancies are part of what separate model-based from model-free behavior. Model-free learning lacks an internal representation of the action-outcome (A-O) contingency, and it has no representation of the outcome itself (liquid vs. food reward, size, flavor, etc.). Instead of using a search and evaluation process to make the decision, model-free decision-making entails choosing the highest “cached value”, derived from a trial and error process of incremental learning. Model-free decision-making is analogous to stimulus-response (S-R) action selection. It is stimulus-bound and outcome blind.

Prospective, reward-related representations in the vStr have been proposed as a possible neural substrate of model-based decision-making [76]. Our data supports this framework. Covert reward decoding in the vStr was greater on deliberative VTE laps than not, and the search process implicated in VTE is characteristically model-based. In line with this idea, a number of studies have shown that the vStr is necessary for certain model-based behaviors. These include reinforcer

devaluation [104-106], identity-based unblocking [92], Pavlovian-Instrumental-Transfer (PIT) [98, 264, 265], and “flexible-approach” behavior [109]. Covert representations of reward could be a general-purpose mechanism by which the vStr contributes to model-based decisions.

Although my data and that in [48] support the role of vStr in prospective reward evaluation, this role has been attributed more often to the OFC [53, 57, 266]. Neural recordings in the OFC show expectancy-related signals of the kind that might be expected in future look-ahead planning [120]. The Schoenbaum lab in particular has provided examples of OFC activity that not only anticipate well-learned outcome associations [267], but also activity that predicts novel, not-yet-experienced outcome values [40, 268, 269]. These latter studies usually (but not always) test whether the rat can make a novel connection between pairs of learned associations ( $A \rightarrow B$ ,  $B \rightarrow C$ ) that have not yet been combined ( $A \rightarrow C$ ). These have variously been called “imagined outcomes”, “novel inference”, and “insight”. These signals, especially “imagined outcomes” [268], bear a close resemblance to the “dynamic evaluation lookahead” [58] signal in vStr seen in my data and that in [48]. As with the vStr, lesion and recording studies in the OFC have implicated it in a number of model-based behaviors [39, 85, 86, 89, 91, 92, 270, 271].

Our data showed covert reward representations in OFC, but only after the moment of TurnAround (Figure 3-13 b). At first glance, this may appear to contradict the literature above. The absence of a forward representation before the moment of choice, predicted by “expectancy theories” of OFC, could be related to the fact that the DD task does not deliver any cues before the rat has committed to his decision. All of the studies mentioned above used cued tasks. Thus, prospective signals in OFC may be expressed differentially based on the presence of overt cues, or in some other manner that is task-dependent. The data presented here shows at the least that the OFC does not universally signal imagined outcomes, and that this process can be accomplished by the vStr.

This is not to say that OFC activity was not contributing to learning or decision-making on the DD task. The increase in OFC decoding was specific to VTE laps, indicative of a role for the

OFC in flexible behavior. The OFC could still be contributing to model-based decision making on the DD task, even in the apparent absence of “imagined outcomes”. In particular, the OFC could be signaling information about the expected value of the *chosen* action, which is necessary for generating the reward prediction error (RPE) signal [2]. Indeed, lesions to the OFC disrupt learning from unexpected outcomes and disrupt reward prediction error signaling in VTA neurons [272, 273]. Given that the reward-prediction error signal is central to both model-free and model-based learning algorithms, the necessity of OFC in generating RPEs is not conclusive evidence for a model-based contribution. However, data in the latter study [272] was most consistent with a computational model in which the OFC signaled expected value, but in addition, OFC input allowed the VTA to differentiate different, but similar states. Thus, the OFC likely transmits expected value information to the VTA, along with state-based information, which by definition is model-based.

This account is consistent with recent empirical work and theoretical proposals arguing that the OFC signals (model-based) information about the current task state [43, 45, 260]. In this formulation, the OFC is still critically involved in forming stimulus-reward associations, and in signaling anticipated reward, but this activity is modulated by task-specific variables that affect reward outcomes, such as the contextual and spatial variables that determine which actions are rewarded [43]. Covert reward-site decoding during VTE (Figure 3-13 b)[120] could represent a chosen value signal, and/or information linking the chosen action with subsequent reward [87, 188]. Further work would be needed to tease apart these different accounts.

The second result from our decoding analyses was that the vStr also preceded the OFC in representing the rats’ impending choice (Figure 3-14). This result is not concordant with the theory that the OFC calculates the economic value of different options before a decision is made [20, 35]; in effect, that the OFC is the locus of value calculation and perhaps also the decision itself. In a study by the Padoa-Schioppa group, chosen value signals emerged earlier in OFC than in other prefrontal structures [274]. Somewhat confusingly, the same group seems to propose that subjective

value signals calculated in the OFC are sent to other structures, where the “goods to action” step takes place [275], putting the locus of action selection outside the OFC. Regardless, the temporal order of activity in OFC and vStr on the DD task does not support this model. Instead, the earlier predictive activity of vStr ensembles suggests a stronger role for vStr than OFC in action-selection, at least in this kind of task (free choice, self-initiated). A recent recording study in primates also found earlier chosen value signals in the vStr as compared to the vmPFC/OFC [51]. See Figure 5-1 for a side by side illustration of these results. These data are consistent with the notion of the vStr as a limbic-motor interface [103], a site where affective and value-related information is translated into a motor plan. It also has intuitive appeal, in that the vStr (and basal ganglia generally) are anatomically/synaptically closer to motor output structures in the brainstem (SNr, superior colliculus, etc.).

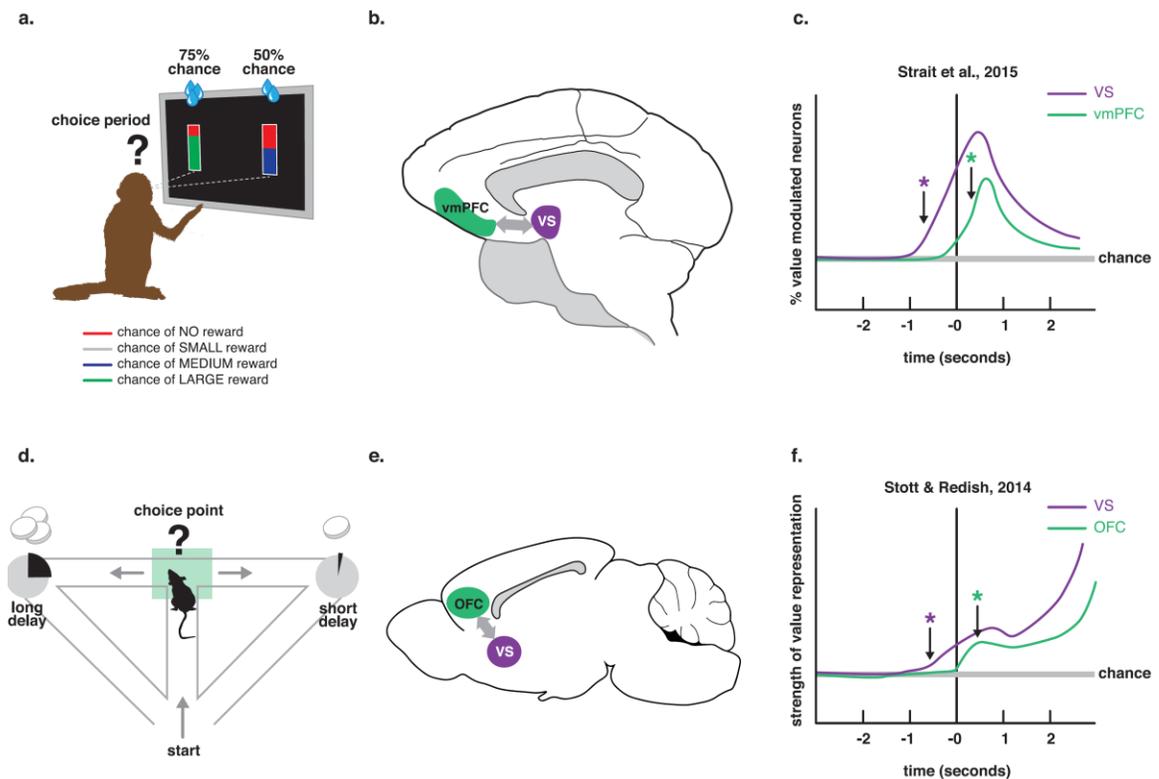


Figure 5-1. **Schematic illustration of parallel neuroeconomic tasks in monkey and rat.** The top panel illustrates the task (a), recording locations (b), and one of the significant results (c), from the

paper by [51]. The bottom panel illustrates the task (**d**), recording locations (**e**), and one of the significant results (**f**), from the paper by [144]. The plot in panel (**c**) is a rendering of Figure 4 C from [51], which measured the percentage of neurons modulated by chosen minus the unchosen value, as a function of time during the trial. The plot in panel (**f**) is a rendering of the data in Figure 3-13 of this manuscript (Figure S6 from [144]). Data was normalized by the mean pFeeder values in the period [-2 -1] sec before TurnAround. Asterisks denote the first time at which the dependent variables reached significance. Note that the ventral striatum (VS) reaches significance before the vmPFC/OFC<sup>18</sup> in both studies. Image credit: Karin Odell. Figure is from Stott & Redish, 2015 [276], freely available for reproduction under the Creative Commons CC0 public domain dedication.

How these chosen value signals are translated into actions is an interesting question for both datasets [51, 144]. Strait et al. found “antagonistic coding” of competing offer values, which they interpret as favoring a mutual inhibition process of action selection [277]. A similar process could be taking place in the vStr on the DD task. This model relies on within-structure processing of competing options. On the DD task, this could entail competition between large-reward preferring neurons and small-reward preferring neurons (Figure 3-7). These neurons would in turn would have to have some way of biasing the motor plan in favor of running to either the large or the small reward arms, respectively. This formulation has received some support in cortex [277, 278]. Decisions between alternatives are thought to result from competition between groups of principal cells that are connected through reciprocal inhibition via interneurons.

Although action-selection models of this nature have been proposed for the striatum [279], they depend critically upon reciprocal inhibition between neurons within the striatum. The main cell type in the striatum—GABAergic medium spiny neurons—have only been shown to inhibit one another in *in vitro* studies [280, 281]. Alternatively, action-selection could take place downstream, in other basal ganglia nuclei (GPi, SNr), or in their outputs (thalamus, motor nuclei).

---

<sup>18</sup> While the authors called their recording site vmPFC, the location of their prefrontal recordings overlap with what primate researchers often label as medial OFC. See Strait et al, 2015 [276], Supplemental Material, for a reconstruction of their recording location.

This is an important outstanding question, as contemporary models of action selection based on temporal difference learning posit that cortical signals are translated into action values within the striatum [61, 77].

To summarize the single unit and decoding data, we found that the OFC and vStr showed very similar activity patterns by some measures, but they also showed distinct differences in the timing of activity, with implications for the roles of OFC and vStr during the decision-making process. The vStr showed earlier activation than the OFC in prospective reward-related and choice-related activity, demonstrating the importance of the vStr in flexible decision-making and action selection, in contrast to top-down theories of cognitive control that emphasize the importance and priority of the neocortex (including OFC) in goal-directed decision-making [3, 8, 20, 231, 242, 282].

The discussion above emphasizes the importance of the vStr on the delay-discounting task. Although the vStr showed temporal precedence over the OFC on key measures of reward-related and choice-related activity, these structures likely work together during value-based decision-making. The OFC and vStr are connected through cortico-striatal connections [123] and through re-entrant cortico-striatal loops. Two recent reports highlight the importance of the orbitofrontal-striatal pathway. Optogenetic stimulation of OFC terminals in the striatum modulated OCD-like compulsive grooming [283, 284]. The temporal order of activity on the DD task suggests that the vStr may be conveying information to the OFC, but this idea was untested. In order to study potential interactions between OFC and vStr, and to better understand local information processing, we applied local field potential analyses to characterize oscillatory dynamics within and between these structures.

#### Fundamental frequencies in OFC and vStr on the DD task

Power spectral density plots (Figure 4-1), spectrograms (Figure 4-2), and self-coherence plots (Figure 4-3) all showed significant power in the beta and gamma range for OFC and vStr.

Additionally, vStr showed prominent gamma80 power, as previously reported [119]. Although not a great deal is known about oscillatory activity in the vStr (as compared, say, to the hippocampus), or particularly in the OFC, this basic spectral profile matches other reports in the literature. Other authors have found significant gamma power in the OFC at around 50-60 Hz [150, 151]. These same studies, and an additional one [221], report significant theta power in the OFC as well. This was not apparent in my analyses, but may be due to task differences, or the time epoch of spectral calculation. Theta power in those studies was high during odor sampling, as was also found in the olfactory bulb [285].

Similar to the OFC, other authors have consistently reported a significant band of gamma power in the vStr, at roughly 50 Hz+ [119, 147-149], with two of these reporting the broad gamma80 shoulder seen in our data [119, 148]. Theta activity was also seen in the PSDs in some of these studies [147, 149], but not others [119, 148], likely owing to the low amplitude of the theta signal (see [286] Figure 9). Significant beta peaks (15-30 Hz) were also seen in each of these reports, although this frequency band has received less attention (but see [211, 212]). Overall, the most prominent frequencies in the OFC and vStr were in the gamma range, which has been specifically implicated in inter-regional communication [223, 287, 288].

An important consideration when studying LFP signals is the issue of volume conduction. The LFP reflects aggregate activity over a large population of cells, predominantly (though not exclusively) synaptic potentials [289]. Electrical activity travels through the extracellular medium, and therefore, local field potentials can reflect electrical field changes arising from other brain regions [290]. This mode of transmission is called volume conduction. Volume conduction can pose a problem in interpreting LFP signals if one draws conclusions based on the assumption that the signal is locally generated, if in fact it has a distal source. This issue has been raised with regard to the vStr [291] because of its proximity to the piriform cortex, which shows strong gamma50 oscillations [148, 285].

The key question for our purposes is whether gamma oscillations in the vStr (and OFC, because it too is adjacent to piriform) are locally relevant. Even if gamma oscillations in the OFC and vStr have a non-local component, the LFP is still relevant if it relates systematically to local spiking activity. Other studies are very clear on this point. A significant fraction of cells in the OFC [150] and vStr [119, 147, 149] show phase locking at gamma frequency. Intracellular recordings from parvalbumin-positive fast-spiking interneurons (FSIs) revealed that these cells have an intrinsic subthreshold oscillation at ~50 Hz [292], and intracellular [293] and extracellular [294] recordings show that FSIs fire preferentially at gamma frequency in response to synaptic input, consistent with a large body of evidence that FSIs coordinate local gamma rhythms, as has been found in other brain regions [295-300]. Additionally, grid recordings across the vStr showed gradients in gamma phase that were inconsistent with volume conduction [301]. Thus, there is substantial evidence that gamma oscillations in the vStr at least may be generated locally through a combination of intrinsic and network interactions. Although volume conduction cannot be ruled out entirely, gamma oscillations in the OFC and vStr are relevant to local information processing in neurons and show distinct correlates to decision-making behavior [128, 129].

#### Differential dynamics in gamma-band activity

To evaluate gamma-band dynamics, we constructed PETHs of gamma50 and gamma80 power, aligned to behaviorally relevant time points. At the choice point, gamma50 power in vStr increased transiently after the point of TurnAround, while gamma80 in the vStr and gamma power in the OFC did not show a clear change at the time of decision (Figure 4-5). This increase in vStr gamma50 power replicates the increase in gamma50 seen in [119] at the time of reversal (when the rat reverses direction on “wrong way” laps), and is most likely related to the increase in gamma50 that is seen throughout the striatum prior to movement onset [152]. This increase was not observed in OFC (Figure 4-5 C). Another group has reported (weak) correlations between gamma power and both movement velocity and acceleration (filtering 30-100 Hz). However, regressing out velocity

and acceleration in their analyses did not affect the influence of reward variables on LFP gamma power [149]. Likewise, in the study by [119], speed did not account for gamma power dynamics in general. Thus, gamma50 in the ventral striatal LFP can be correlated with movement onset, as also seen in dorsal striatum [152], but this effect does not explain reward-related changes. Whether or not gamma50 power in either the dorsal or ventral striatum is causally related to movement initiation remains unknown.

Gamma power in the vStr has been consistently implicated in reward processing [128]. On the DD task, gamma50 power rapidly increased after feeder fire, showing two peaks (one after the feeder trigger, and one after reward receipt ~2sec later) (Figure 4-6 A,C). In contrast, gamma80 power transiently dropped after the feeder trigger event and only peaked again some 4 seconds later (Figure 4-6 B,D). Interestingly, gamma power remained high during the adjusting delay, which parallels the sustained increase in firing rate during this period (Figure 3-12). The pattern of gamma modulation seen here closely matches the dynamics seen in [119] (Figure 9). Likewise, [149] reported a transient increase in gamma50 power to reward delivery and a “dramatic” drop in gamma80 power after arrival at the reward site. These contrasting dynamics support the argument for a functional distinction between gamma50 and gamma80 in the vStr [128] and OFC.

As stated above, gamma50 and gamma80 power showed contrasting activity at the time of reward receipt. Specifically, gamma50 power rose when gamma80 power fell. This result matches the spontaneous alternation in gamma power seen in the example spectrograms (Figure 4-4) and the anti-correlation between gamma50 and gamma80 seen in the vStr self-coherence plot (Figure 4-3, B). It also matches the results in [148]. In that study, gamma power showed a switch between frequencies following reward receipt (Figure 4 from that paper). Berke additionally showed that gamma power switched from ~50 Hz to ~80 Hz after delivery of amphetamine or the dopamine agonist apomorphine [148]. These alternating dynamics raise the interesting possibility that distinct frequency bands in the vStr could implement a kind of “switchboard”, allowing the vStr to communicate selectively with different brain areas at different times.

The nucleus accumbens (NAc) receives prominent inputs from the hippocampus, prefrontal cortex, and amygdala [302]. An important open question for vStr function is how these inputs integrate and if there is a selection process to differentially “attend” to particular inputs. According to the switchboard concept, the NAc can switch between different input streams by synchronizing with different brain areas [206], possibly mediated by differential activation of D1 versus D2 dopamine receptors [303]. This same basic idea has also been proposed in terms of action selection and motor output [279]. Thus, frequency-dependent, coherent activity between the vStr and other structures could be indicative of a selective emphasis on different input or output channels. Gamma50 and gamma80 activity could then reflect synchronization with specific brain structures, enabling a transition between different inputs (or outputs).

This proposal has received some empirical support. The vStr shows coherence with other structures at distinct frequencies: hippocampus at theta frequency [148, 286], piriform at 50 Hz [148], mPFC at 80 Hz [148, 226]. On a reward-seeking task, NAc activity was more synchronous with the hippocampus than mPFC during spatial exploration and it was more synchronous with the mPFC than hippocampus during lever pressing [206], as might be expected based on the roles of these two structures. On the Multiple-T maze task, gamma80 power (but not gamma50 power) “ramped up” as rats approached the reward sites [119]. This latter finding links gamma80 activity with reward anticipation/receipt, as in [148], who also showed that sustained dopamine enhancement resulted in a shift from gamma50 to gamma80. Bridging these two results is a recent paper that showed that tonic dopamine transmission in the vStr ramped up as rats navigated down a T-maze toward a goal [230]. Tonic dopamine transmission increased as a function of the spatial and temporal proximity to reward, and also scaled with reward magnitude. Thus, the gamma80 “channel” could be preferentially related to reward in some way (although this does not rule out a role for gamma50). It would be interesting to see if stimulation of DA at a fast timescale increases gamma80 power in vStr and if it increases coherence between the vStr and mPFC.

### Functional coupling between OFC and vStr on the DD task

While there has been some work investigated interactions between the mPFC (areas PL and IL) and vStr in the rodent [206, 304], little is known about communication between the OFC and vStr, despite the fact that these structures are also directly connected and both involved in reward-guided behavior. Do the OFC and vStr also show functional coupling? At the whole session level, coherence between OFC and vStr was high, and showed distinct peaks at delta, beta, and both gamma bands (Figure 4-7). Gamma50 showed the strongest coherence. We focused our coherence analyses on gamma band activity, as this has been linked to decision-making in both structures [128, 129], and is especially implicated in long-range communication between brain structures [130, 288].

At the choice point, gamma50 showed a peak in coherence after the time of TurnAround on VTE laps, but this was not present in the gamma80 range (Figure 4-8). This peak may be related to the increase in vStr power seen after TurnAround (Figure 4-5), which is seen just prior to movement initiation [152]. Rats on T-maze tasks such as the DD tasks show movement initiation events both at the choice point during VTE and when they leave the feeder site. The relationship between these two behaviors and gamma50 bursts has not been directly compared. While gamma50 after TurnAround may reflect movement initiation, it differs from movement onset at the feeder site in that the former is not directly preceded by reward receipt. Thus, the functional significance of gamma50 at the choice point is not yet known.

How does OFC-vStr coherence change around the time of reward receipt, when both structures are particularly active? Coherence at the time of feeder Zone entry (same moment at Choice Point Exit) followed roughly the same time course as gamma power, similarly aligned (Figure 4-6 vs. Figure 4-9). In both gamma bands, there was a multiphasic relationship. Gamma50 coherence showed a double peak, likely corresponding to the auditory cue that signaled zone entry, and then subsequent reward receipt (Figure 4-9 A). Gamma80 coherence, in contrast, showed a

strong, phasic decrease after zone entry (Figure 4-9 B). Coherence around the time of reward receipt was not substantially different between VTE and non-VTE laps. These data indicate that there is significant coherence between OFC and vStr, especially in the gamma50 band (higher absolute values), during reward anticipation and receipt. These dynamics differed between gamma50 and gamma80 in a way that paralleled the changes in power in those structures. This indicates a general (broadband) level of synchrony in the gamma range between OFC and vStr, in contrast to the narrower range of synchrony (gamma80) seen between vStr and mPFC [148, 226].

Coherence at both gamma50 and gamma80 would seem to argue against the switchboard concept of vStr processing. In terms of frequency matching, there would be interference (or competition) between OFC and mPFC, unless coherence between these two pairs (OFC-vStr, mPFC-vStr) occurred at different times, or if there were some additional means of filtering. However, different cortical sites have distinct topographical representations in the striatum [126, 305]. While there is some degree of overlap in cortical innervation (the exact mapping depends on the techniques used and the species under study), there is also substantial segregation. Current thinking holds that different cortico-basal ganglia loops process information in parallel, leaving room for mPFC and OFC to interact with the striatum in parallel as well.

#### Directed connectivity between OFC and vStr on the DD task

As we have seen, the OFC and vStr show prominent coherence, especially in the gamma range, which is modulated by task events. This naturally leads to the question of which structure is influencing the other, and when. Does the vStr inherit value signals from the OFC, as might be suggested by OFC-centric models of valuation and decision-making [35, 57]? Or does the vStr calculate value first and pass it along to the OFC, as suggested by the timing of valuation signals in these two areas (decoding data (Chapter 3); [276])? Recent advances in analytical techniques allow us to address this question [163, 164, 171]. We applied independent, but complementary

analyses to the LFP data from OFC and vStr in order to investigate directed connectivity between these structures.

Granger causality analyses revealed that a much larger fraction of sessions had greater GC influence in the vStr→OFC direction than in the OFC→vStr direction (Figure 4-10 A). In other words, local field potential values in the vStr better predicted upcoming local field potential values in the OFC than vice versa. When we decomposed the time-domain GC values according to their spectral components, we observed greater GC values for vStr→OFC than vice versa, most prominently at 50 Hz (Figure 4-10 B). This “molar analysis” also held true for behaviorally relevant task epochs. Time-domain GC values were greater from vStr to OFC than in the opposite direction, around the time of reward delivery and at the choice point (Figure 4-11 A,B, respectively). Likewise, cross-correlation analyses revealed that LFP signals in the OFC lagged those in the vStr in three out of the four frequency bands tested, with the gamma band being most significant (absolute p values) (Figure 4-12). The cross-correlation analyses support the Granger causality results, demonstrating that vStr leads OFC at the local field potential level. Thus, LFP signals in the vStr showed temporal *precedence* with respect to OFC, and the vStr was more *predictive* of subsequent OFC activity. We can conclude that the direction of information transmission on the DD task is chiefly in the vStr to OFC direction, and that this influence is most concentrated in the 50 Hz range.

What do these results mean? Put succinctly, they reverse (literally and figuratively) the direction of information flow. In most box and arrow diagrams of decision-making, the cortex sends an arrow in the “feedforward” direction to the basal ganglia. The arrow from the basal ganglia to the cortex is considered secondarily, regarded as “feedback”, and is thought to implement a gating or modulatory effect on cortical firing. The timing of activity and the direction of information transmission on the DD task was in the opposite direction: from the basal ganglia to the neocortex. With that said, the data from our experiments do not obviate a feedforward role for the frontal cortex, nor do they invalidate the literature showing that various value-related computations can be

executed in the cortex. However, we argue that the results presented in this thesis make evident that in certain situations the arrow can run in the opposite direction.

How do the LFP results relate to the results in Chapter 3? First and foremost, the LFP data are congruent with the Bayesian decoding analyses. In those analyses, the vStr showed covert reward signaling earlier than the OFC on VTE trials, and the vStr showed choice-predictive activity earlier than the OFC generally (on VTE and non-VTE laps). Those data established that the vStr calculated, or at least had access to, value and choice-related information before the OFC on a neuroeconomic decision-making task. This in itself is a novel result, and noteworthy because it runs contrary to leading theories that maintain that it is the cortex which calculates value and sends that information “downstream” to either bias or execute a motor plan [275]. This point is all the more noteworthy for goal-directed decision-making, which is thought to take place “on the fly”, and involves explicit knowledge about the expected reward. In the decision-making literature, the striatum (including the vStr) is often restricted to implementing slow, trial and error learning (i.e. feedback) processes [77], while the cortex is thought to be uniquely capable of flexible, “online” behavioral control [231].

The local field potential data extends the decoding results by showing that the vStr exerts a stronger causal influence over the OFC than vice versa—particularly during VTE (Figure 4-11 B)—suggesting that the vStr may be sending evaluation-related information to the OFC (contrary to the standard model). Extending this line of thought, the OFC may be translating value-related information from the vStr into *chosen value* signals [188], or perhaps *state value* signals necessary for calculating reward prediction errors [272, 273]. Further work is needed to understand what information is being communicated from the vStr to the OFC. The data presented in this thesis suggests that this is a fruitful line of inquiry.

## Bibliography

1. Stephens, D.W. and J.R. Krebs, *Foraging theory*. 1987, Princeton, N.J.: Princeton University Press.
2. Sutton, R.S. and A.G. Barto, *Reinforcement learning : an introduction*. Adaptive computation and machine learning. 1998, Cambridge, Mass.: MIT Press. xviii, 322 p.
3. Rangel, A. and T. Hare, *Neural computations associated with goal-directed choice*. *Curr Opin Neurobiol*, 2010. **20**(2): p. 262-70.
4. von Neumann, J. and O. Morgenstern, *Theory of Games and Economic Behavior*. 1963, Princeton, NJ: Princeton University Press.
5. Glimcher, P.W., *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*. *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics*, 2003: p. 1-375.
6. Gul, F., and Pessendorfer, *The case for mindless economics*. , in *Foundations of positive and normative economics*. , A.S. Andrew Caplin, Editor. 2008, Oxford University Press: New York. p. 3-39.
7. Glimcher, P.W. and A. Rustichini, *Neuroeconomics: the consilience of brain and decision*. *Science*, 2004. **306**(5695): p. 447-52.
8. Levy, D.J. and P.W. Glimcher, *The root of all value: a neural common currency for choice*. *Current Opinion in Neurobiology*, 2012. **22**(6): p. 1027-1038.
9. Glimcher, P.W. and E. Fehr, *Neuroeconomics : decision making and the brain*. Second edition. ed. 2014, Amsterdam Boston: Academic Press. xxviii, 577 pages.
10. Breiter, H.C., et al., *Functional imaging of neural responses to expectancy and experience of monetary gains and losses*. *Neuron*, 2001. **30**(2): p. 619-39.
11. Knutson, B., et al., *Anticipation of increasing monetary reward selectively recruits nucleus accumbens*. *J Neurosci*, 2001. **21**(16): p. Rc159.
12. Knutson, B., et al., *Distributed neural representation of expected value*. *J Neurosci*, 2005. **25**(19): p. 4806-12.
13. O'Doherty, J.P., *Reward representations and reward-related learning in the human brain: insights from neuroimaging*. *Curr Opin Neurobiol*, 2004. **14**(6): p. 769-76.
14. Kable, J.W. and P.W. Glimcher, *The neural correlates of subjective value during intertemporal choice*. *Nat Neurosci*, 2007. **10**(12): p. 1625-33.
15. Levy, I., et al., *Neural representation of subjective value under risk and ambiguity*. *J Neurophysiol*, 2010. **103**(2): p. 1036-47.
16. Diekhof, E.K., et al., *The role of the human ventral striatum and the medial orbitofrontal cortex in the representation of reward magnitude - an activation likelihood estimation meta-analysis of neuroimaging studies of passive reward expectancy and outcome processing*. *Neuropsychologia*, 2012. **50**(7): p. 1252-66.
17. Padoa-Schioppa, C. and J.A. Assad, *Neurons in the orbitofrontal cortex encode economic value*. *Nature*, 2006. **441**(7090): p. 223-6.
18. Padoa-Schioppa, C. and J.A. Assad, *The representation of economic value in the orbitofrontal cortex is invariant for changes of menu*. *Nat Neurosci*, 2008. **11**(1): p. 95-102.
19. Padoa-Schioppa, C., *Range-adapting representation of economic value in the orbitofrontal cortex*. *J Neurosci*, 2009. **29**(44): p. 14004-14.
20. Padoa-Schioppa, C. and X. Cai, *The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives*. *Ann N Y Acad Sci*, 2011. **1239**: p. 130-7.

21. Kennerley, S.W. and J.D. Wallis, *Encoding of reward and space during a working memory task in the orbitofrontal cortex and anterior cingulate sulcus*. J Neurophysiol, 2009. **102**(6): p. 3352-64.
22. Kennerley, S.W. and J.D. Wallis, *Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables*. Eur J Neurosci, 2009. **29**(10): p. 2061-73.
23. O'Neill, M. and W. Schultz, *Economic risk coding by single neurons in the orbitofrontal cortex*. J Physiol Paris, 2015. **109**(1-3): p. 70-7.
24. Tremblay, L. and W. Schultz, *Relative reward preference in primate orbitofrontal cortex*. Nature, 1999. **398**(6729): p. 704-8.
25. Roesch, M.R. and C.R. Olson, *Neuronal activity related to reward value and motivation in primate frontal cortex*. Science, 2004. **304**(5668): p. 307-10.
26. Critchley, H.D. and E.T. Rolls, *Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex*. J Neurophysiol, 1996. **75**(4): p. 1673-86.
27. Grattan, L.E. and P.W. Glimcher, *Absence of spatial tuning in the orbitofrontal cortex*. PLoS One, 2014. **9**(11): p. e112750.
28. O'Neill, M. and W. Schultz, *Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value*. Neuron, 2010. **68**(4): p. 789-800.
29. Morrison, S.E. and C.D. Salzman, *The convergence of information about rewarding and aversive stimuli in single neurons*. J Neurosci, 2009. **29**(37): p. 11471-83.
30. Fehr, E. and A. Rangel, *Neuroeconomic Foundations of Economic Choice-Recent Advances*. Journal of Economic Perspectives, 2011. **25**(4): p. 3-30.
31. Hare, T.A., C.F. Camerer, and A. Rangel, *Self-control in decision-making involves modulation of the vmPFC valuation system*. Science, 2009. **324**(5927): p. 646-8.
32. Bartra, O., J.T. McGuire, and J.W. Kable, *The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value*. Neuroimage, 2013. **76**: p. 412-27.
33. Wallis, J.D. and S.W. Kennerley, *Heterogeneous reward signals in prefrontal cortex*. Curr Opin Neurobiol, 2010. **20**(2): p. 191-8.
34. Grabenhorst, F. and E.T. Rolls, *Value, pleasure and choice in the ventral prefrontal cortex*. Trends Cogn Sci, 2011. **15**(2): p. 56-67.
35. Padoa-Schioppa, C., *Neurobiology of economic choice: a good-based model*. Annu Rev Neurosci, 2011. **34**: p. 333-59.
36. Schoenbaum, G., et al., *Does the orbitofrontal cortex signal value?* Ann N Y Acad Sci, 2011. **1239**: p. 87-99.
37. Roesch, M.R., A.R. Taylor, and G. Schoenbaum, *Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation*. Neuron, 2006. **51**(4): p. 509-20.
38. Feierstein, C.E., et al., *Representation of spatial goals in rat orbitofrontal cortex*. Neuron, 2006. **51**(4): p. 495-507.
39. McDannald, M.A., et al., *Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking*. Elife, 2014. **3**: p. e02653.
40. Stalnaker, T.A., et al., *Orbitofrontal neurons infer the value and identity of predicted outcomes*. Nat Commun, 2014. **5**: p. 3926.
41. Kepecs, A., et al., *Neural correlates, computation and behavioural impact of decision confidence*. Nature, 2008. **455**(7210): p. 227-31.

42. Sul, J.H., et al., *Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making*. *Neuron*, 2010. **66**(3): p. 449-60.
43. Farovik, A., et al., *Orbitofrontal Cortex Encodes Memories within Value-Based Schemas and Represents Contexts That Guide Memory Retrieval*. *J Neurosci*, 2015. **35**(21): p. 8333-44.
44. Steiner, A.P. and A.D. Redish, *Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task*. *Nat Neurosci*, 2014. **17**(7): p. 995-1002.
45. Wilson, R.C., et al., *Orbitofrontal cortex as a cognitive map of task space*. *Neuron*, 2014. **81**(2): p. 267-79.
46. Lavoie, A.M. and S.J. Mizumori, *Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of rats*. *Brain Res*, 1994. **638**(1-2): p. 157-68.
47. Lansink, C.S., et al., *Fast-spiking interneurons of the rat ventral striatum: temporal coordination of activity with principal cells and responsiveness to reward*. *Eur J Neurosci*, 2010. **32**(3): p. 494-508.
48. van der Meer, M.A. and A.D. Redish, *Covert Expectation-of-Reward in Rat Ventral Striatum at Decision Points*. *Front Integr Neurosci*, 2009. **3**: p. 1.
49. Roesch, M.R., et al., *Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards*. *J Neurosci*, 2009. **29**(42): p. 13365-76.
50. German, P.W. and H.L. Fields, *Rat nucleus accumbens neurons persistently encode locations associated with morphine reward*. *J Neurophysiol*, 2007. **97**(3): p. 2094-106.
51. Strait, C.E., Sleezer B.J., Hayden, B.Y., *Signatures of value comparison in ventral striatum neurons*. *PLOS Biology*, 2015.
52. Schultz, W., L. Tremblay, and J.R. Hollerman, *Reward processing in primate orbitofrontal cortex and basal ganglia*. *Cereb Cortex*, 2000. **10**(3): p. 272-84.
53. Schoenbaum, G. and M. Roesch, *Orbitofrontal cortex, associative learning, and expectancies*. *Neuron*, 2005. **47**(5): p. 633-6.
54. Carelli, R.M., *The nucleus accumbens and reward: neurophysiological investigations in behaving animals*. *Behav Cogn Neurosci Rev*, 2002. **1**(4): p. 281-96.
55. Rolls, E.T., H.D. Critchley, and A. Treves, *Representation of olfactory information in the primate orbitofrontal cortex*. *J Neurophysiol*, 1996. **75**(5): p. 1982-96.
56. Rolls, E.T., *The orbitofrontal cortex and reward*. *Cereb Cortex*, 2000. **10**(3): p. 284-94.
57. Rudebeck, P.H. and E.A. Murray, *The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes*. *Neuron*, 2014. **84**(6): p. 1143-56.
58. van der Meer, M.A. and A.D. Redish, *Expectancies in decision making, reinforcement learning, and ventral striatum*. *Front Neurosci*, 2010. **4**: p. 6.
59. Balleine, B.W. and A. Dickinson, *Goal-directed instrumental action: contingency and incentive learning and their cortical substrates*. *Neuropharmacology*, 1998. **37**(4-5): p. 407-19.
60. Doya, K., *What are the computations of the cerebellum, the basal ganglia and the cerebral cortex?* *Neural Netw*, 1999. **12**(7-8): p. 961-974.
61. Daw, N.D., Y. Niv, and P. Dayan, *Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control*. *Nat Neurosci*, 2005. **8**(12): p. 1704-11.
62. Yin, H.H. and B.J. Knowlton, *The role of the basal ganglia in habit formation*. *Nat Rev Neurosci*, 2006. **7**(6): p. 464-76.

63. Rangel, A., C. Camerer, and P.R. Montague, *A framework for studying the neurobiology of value-based decision making*. Nat Rev Neurosci, 2008. **9**(7): p. 545-56.
64. Redish, A.D., *The mind within the brain : how we make decisions and how those decisions go wrong*. 2013, Oxford: Oxford University Press. xii, 377 pages.
65. Redish, A.D., S. Jensen, and A. Johnson, *A unified framework for addiction: vulnerabilities in the decision process*. Behav Brain Sci, 2008. **31**(4): p. 415-37; discussion 437-87.
66. Balleine, B.W. and J.P. O'Doherty, *Human and rodent homologues in action control: corticostriatal determinants of goal-directed and habitual action*. Neuropsychopharmacology, 2010. **35**(1): p. 48-69.
67. Mackintosh, N.J., *Animal learning and cognition*. Handbook of perception and cognition (2nd ed ). 1994, San Diego: Academic Press. xviii, 379 p.
68. Graybiel, A.M., *Habits, rituals, and the evaluative brain*. Annu Rev Neurosci, 2008. **31**: p. 359-87.
69. van der Meer, M., Z. Kurth-Nelson, and A.D. Redish, *Information processing in decision-making systems*. Neuroscientist, 2012. **18**(4): p. 342-59.
70. Lee, S.W., S. Shimojo, and J.P. O'Doherty, *Neural computations underlying arbitration between model-based and model-free learning*. Neuron, 2014. **81**(3): p. 687-99.
71. Daw, N.D., et al., *Cortical substrates for exploratory decisions in humans*. Nature, 2006. **441**(7095): p. 876-9.
72. Johnson, A., et al., *The hippocampus and exploration: dynamically evolving behavior and neural representations*. Front Hum Neurosci, 2012. **6**: p. 216.
73. O'Keefe, J. and L. Nadel, *The hippocampus as a cognitive map*. 1978, Oxford: Oxford University Press. xiv, 570 p.
74. Redish, A.D., *Beyond the cognitive map : from place cells to episodic memory*. 1999, Cambridge, Mass.: MIT Press. xviii, 420 p.
75. Cohen, J.D., S.M. McClure, and A.J. Yu, *Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration*. Philos Trans R Soc Lond B Biol Sci, 2007. **362**(1481): p. 933-42.
76. Johnson, A., M.A. van der Meer, and A.D. Redish, *Integrating hippocampus and striatum in decision-making*. Curr Opin Neurobiol, 2007. **17**(6): p. 692-7.
77. Takahashi, Y., G. Schoenbaum, and Y. Niv, *Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model*. Front Neurosci, 2008. **2**(1): p. 86-99.
78. Packard, M.G. and J.L. McGaugh, *Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning*. Neurobiol Learn Mem, 1996. **65**(1): p. 65-72.
79. Jog, M.S., et al., *Building neural representations of habits*. Science, 1999. **286**(5445): p. 1745-9.
80. Tricomi, E., B.W. Balleine, and J.P. O'Doherty, *A specific role for posterior dorsolateral striatum in human habit learning*. Eur J Neurosci, 2009. **29**(11): p. 2225-32.
81. Barnes, T.D., et al., *Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories*. Nature, 2005. **437**(7062): p. 1158-61.
82. Smith, K.S. and A.M. Graybiel, *A dual operator view of habitual behavior reflecting cortical and striatal dynamics*. Neuron, 2013. **79**(2): p. 361-74.
83. Regier, P.S., S. Amemiyia, and A.D. Redish, *Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task*. J Neurophysiol, 2015: p. jn 00189 2015.

84. Jin, X. and R.M. Costa, *Start/stop signals emerge in nigrostriatal circuits during sequence learning*. Nature, 2010. **466**(7305): p. 457-62.
85. Jones, J.L., et al., *Orbitofrontal cortex supports behavior and learning using inferred but not cached values*. Science, 2012. **338**(6109): p. 953-6.
86. McDannald, M.A., et al., *Model-based learning and the contribution of the orbitofrontal cortex to the model-free world*. Eur J Neurosci, 2012. **35**(7): p. 991-6.
87. Walton, M.E., et al., *Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning*. Neuron, 2010. **65**(6): p. 927-39.
88. Ostlund, S.B. and B.W. Balleine, *Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning*. J Neurosci, 2007. **27**(18): p. 4819-25.
89. Pickens, C.L., et al., *Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task*. J Neurosci, 2003. **23**(35): p. 11078-84.
90. Gallagher, M., R.W. McMahan, and G. Schoenbaum, *Orbitofrontal cortex and representation of incentive value in associative learning*. J Neurosci, 1999. **19**(15): p. 6610-4.
91. Izquierdo, A., R.K. Suda, and E.A. Murray, *Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency*. J Neurosci, 2004. **24**(34): p. 7540-8.
92. McDannald, M.A., et al., *Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning*. J Neurosci, 2011. **31**(7): p. 2700-5.
93. Padoa-Schioppa, C., *Orbitofrontal cortex and the computation of economic value*. Ann N Y Acad Sci, 2007. **1121**: p. 232-53.
94. Yin, H.H., et al., *The role of the dorsomedial striatum in instrumental conditioning*. Eur J Neurosci, 2005. **22**(2): p. 513-23.
95. Salamone, J.D. and M. Correa, *Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine*. Behav Brain Res, 2002. **137**(1-2): p. 3-25.
96. Salamone, J.D., et al., *Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits*. Psychopharmacology (Berl), 2007. **191**(3): p. 461-82.
97. Berridge, K.C., T.E. Robinson, and J.W. Aldridge, *Dissecting components of reward: 'liking', 'wanting', and learning*. Curr Opin Pharmacol, 2009. **9**(1): p. 65-73.
98. Corbit, L.H. and B.W. Balleine, *The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell*. J Neurosci, 2011. **31**(33): p. 11786-94.
99. Setlow, B., P.C. Holland, and M. Gallagher, *Disconnection of the basolateral amygdala complex and nucleus accumbens impairs appetitive pavlovian second-order conditioned responses*. Behav Neurosci, 2002. **116**(2): p. 267-75.
100. Holland, P.C. and R.A. Rescorla, *The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning*. J Exp Psychol Anim Behav Process, 1975. **1**(4): p. 355-63.
101. Everitt, B.J., et al., *The basolateral amygdala-ventral striatal system and conditioned place preference: further evidence of limbic-striatal interactions underlying reward-related processes*. Neuroscience, 1991. **42**(1): p. 1-18.
102. Kalivas, P.W., B.A. Sorg, and M.S. Hooks, *The pharmacology and neural circuitry of sensitization to psychostimulants*. Behav Pharmacol, 1993. **4**(4): p. 315-334.

103. Mogenson, G.J., D.L. Jones, and C.Y. Yim, *From motivation to action: functional interface between the limbic system and the motor system*. Prog Neurobiol, 1980. **14**(2-3): p. 69-97.
104. Lex, B. and W. Hauber, *The role of nucleus accumbens dopamine in outcome encoding in instrumental and Pavlovian conditioning*. Neurobiol Learn Mem, 2010. **93**(2): p. 283-90.
105. Singh, T., et al., *Nucleus Accumbens Core and Shell are Necessary for Reinforcer Devaluation Effects on Pavlovian Conditioned Responding*. Front Integr Neurosci, 2010. **4**: p. 126.
106. Corbit, L.H., J.L. Muir, and B.W. Balleine, *The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell*. J Neurosci, 2001. **21**(9): p. 3251-60.
107. McDannald, M.A., et al., *Learning theory: a driving force in understanding orbitofrontal function*. Neurobiol Learn Mem, 2014. **108**: p. 22-7.
108. Holland, P.C., *Unblocking in Pavlovian appetitive conditioning*. J Exp Psychol Anim Behav Process, 1984. **10**(4): p. 476-97.
109. Nicola, S.M., *The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior*. J Neurosci, 2010. **30**(49): p. 16585-600.
110. du Hoffmann, J. and S.M. Nicola, *Dopamine invigorates reward seeking by promoting cue-evoked excitation in the nucleus accumbens*. J Neurosci, 2014. **34**(43): p. 14349-64.
111. Daw, N.D., et al., *Model-based influences on humans' choices and striatal prediction errors*. Neuron, 2011. **69**(6): p. 1204-15.
112. Papale, A.E., et al., *Interactions between deliberation and delay-discounting in rats*. Cogn Affect Behav Neurosci, 2012. **12**(3): p. 513-26.
113. Tolman, E.C., *Cognitive maps in rats and men*. Psychol Rev, 1948. **55**(4): p. 189-208.
114. Johnson, A. and A.D. Redish, *Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point*. J Neurosci, 2007. **27**(45): p. 12176-89.
115. Tolman, E.C., B.F. Ritchie, and D. Kalish, *Studies in spatial learning: Orientation and the short-cut*. J Exp Psychol, 1946. **36**: p. 13-24.
116. Tolman, E.C., *The determiners of behavior at a choice point*. Psychological Review, 1938. **45**(1): p. 1-41.
117. Doll, B.B., et al., *Model-based choices involve prospective neural activity*. Nat Neurosci, 2015. **18**(5): p. 767-72.
118. Bornstein, A.M. and N.D. Daw, *Cortical and hippocampal correlates of deliberation during model-based decisions for rewards in humans*. PLoS Comput Biol, 2013. **9**(12): p. e1003387.
119. van der Meer, M.A. and A.D. Redish, *Low and High Gamma Oscillations in Rat Ventral Striatum have Distinct Relationships to Behavior, Reward, and Spiking Activity on a Learned Spatial Decision Task*. Front Integr Neurosci, 2009. **3**: p. 9.
120. Steiner, A.P. and A.D. Redish, *The road not taken: neural correlates of decision making in orbitofrontal cortex*. Front Neurosci, 2012. **6**: p. 131.
121. Haber, S.N. and B. Knutson, *The reward circuit: linking primate anatomy and human imaging*. Neuropsychopharmacology, 2010. **35**(1): p. 4-26.
122. Voorn, P., et al., *Putting a spin on the dorsal-ventral divide of the striatum*. Trends Neurosci, 2004. **27**(8): p. 468-74.
123. Haber, S.N., et al., *The orbital and medial prefrontal circuit through the primate basal ganglia*. J Neurosci, 1995. **15**(7 Pt 1): p. 4851-67.

124. Brog, J.S., et al., *The patterns of afferent innervation of the core and shell in the "accumbens" part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluoro-gold*. J Comp Neurol, 1993. **338**(2): p. 255-78.
125. Schilman, E.A., et al., *The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex*. Neurosci Lett, 2008. **432**(1): p. 40-5.
126. Maily, P., et al., *The rat prefrontostriatal system analyzed in 3D: evidence for multiple interacting functional units*. J Neurosci, 2013. **33**(13): p. 5718-27.
127. Alexander, G.E. and M.D. Crutcher, *Functional architecture of basal ganglia circuits: neural substrates of parallel processing*. Trends Neurosci, 1990. **13**(7): p. 266-71.
128. van der Meer, M.A., et al., *Integrating early results on ventral striatal gamma oscillations in the rat*. Front Neurosci, 2010. **4**: p. 300.
129. Pennartz, C.M., M. van Wingerden, and M. Vinck, *Population coding and neural rhythmicity in the orbitofrontal cortex*. Ann N Y Acad Sci, 2011. **1239**: p. 149-61.
130. Harris, A.Z. and J.A. Gordon, *Long-Range Neural Synchrony in Behavior*. Annu Rev Neurosci, 2015. **38**: p. 171-94.
131. Winstanley, C.A., et al., *Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice*. J Neurosci, 2004. **24**(20): p. 4718-22.
132. Zeeb, F.D., S.B. Floresco, and C.A. Winstanley, *Contributions of the orbitofrontal cortex to impulsive choice: interactions with basal levels of impulsivity, dopamine signalling, and reward-related cues*. Psychopharmacology (Berl), 2010. **211**(1): p. 87-98.
133. Mobini, S., et al., *Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement*. Psychopharmacology (Berl), 2002. **160**(3): p. 290-8.
134. Rudebeck, P.H., et al., *Separate neural pathways process different decision costs*. Nat Neurosci, 2006. **9**(9): p. 1161-8.
135. Valencia-Torres, L., et al., *Nucleus accumbens and delay discounting in rats: evidence from a new quantitative protocol for analysing inter-temporal choice*. Psychopharmacology (Berl), 2012. **219**(2): p. 271-83.
136. Cardinal, R.N. and T.H. Cheung, *Nucleus accumbens core lesions retard instrumental learning and performance with delayed reinforcement in the rat*. BMC Neurosci, 2005. **6**: p. 9.
137. Cardinal, R.N. and N.J. Howes, *Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats*. BMC Neurosci, 2005. **6**: p. 37.
138. Cardinal, R.N., et al., *Impulsive choice induced in rats by lesions of the nucleus accumbens core*. Science, 2001. **292**(5526): p. 2499-501.
139. Dalley, J.W., et al., *Neurobehavioral mechanisms of impulsivity: fronto-striatal systems and functional neurochemistry*. Pharmacol Biochem Behav, 2008. **90**(2): p. 250-60.
140. da Costa Araujo, S., et al., *Choice between reinforcer delays versus choice between reinforcer magnitudes: differential Fos expression in the orbital prefrontal cortex and nucleus accumbens core*. Behav Brain Res, 2010. **213**(2): p. 269-77.
141. Bezzina, G., et al., *Effect of disconnecting the orbital prefrontal cortex from the nucleus accumbens core on inter-temporal choice behaviour: a quantitative analysis*. Behav Brain Res, 2008. **191**(2): p. 272-9.
142. Cooch, N.K., et al., *Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons*. Nat Commun, 2015. **6**: p. 7195.
143. Cardinal, R.N., *Neural systems implicated in delayed and probabilistic reinforcement*. Neural Netw, 2006. **19**(8): p. 1277-301.

144. Stott, J.J. and A.D. Redish, *A functional difference in information processing between orbitofrontal cortex and ventral striatum during decision-making behaviour*. *Philos Trans R Soc Lond B Biol Sci*, 2014. **369**(1655).
145. Paxinos, G. and C. Watson, *The rat brain in stereotaxic coordinates*. 4th ed. 1998, San Diego: Academic Press.
146. Kadir, S.N., D.F. Goodman, and K.D. Harris, *High-dimensional cluster analysis with the masked EM algorithm*. *Neural Comput*, 2014. **26**(11): p. 2379-94.
147. Berke, J.D., et al., *Oscillatory entrainment of striatal neurons in freely moving rats*. *Neuron*, 2004. **43**(6): p. 883-96.
148. Berke, J.D., *Fast oscillations in cortical-striatal networks switch frequency following rewarding events and stimulant drugs*. *Eur J Neurosci*, 2009. **30**(5): p. 848-59.
149. Kalenscher, T., et al., *Reward-associated gamma oscillations in ventral striatum are regionally differentiated and modulate local firing activity*. *J Neurophysiol*, 2010. **103**(3): p. 1658-72.
150. van Wingerden, M., et al., *Learning-associated gamma-band phase-locking of action-outcome selective neurons in orbitofrontal cortex*. *J Neurosci*, 2010. **30**(30): p. 10025-38.
151. van Wingerden, M., et al., *Phase-amplitude coupling in rat orbitofrontal cortex discriminates between correct and incorrect decisions during associative learning*. *J Neurosci*, 2014. **34**(2): p. 493-505.
152. Masimore, B., et al., *Transient striatal gamma local field potentials signal movement initiation in rats*. *Neuroreport*, 2005. **16**(18): p. 2021-4.
153. Hart, W.E., et al., *Measurement and classification of retinal vascular tortuosity*. *Int J Med Inform*, 1999. **53**(2-3): p. 239-52.
154. Janabi-Sharifi, F., V. Hayward, and C.S.J. Chen, *Discrete-time adaptive windowing for velocity estimation*. *Ieee Transactions on Control Systems Technology*, 2000. **8**(6): p. 1003-1009.
155. Zhang, K., et al., *Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells*. *J Neurophysiol*, 1998. **79**(2): p. 1017-44.
156. Davidson, T.J., F. Kloosterman, and M.A. Wilson, *Hippocampal replay of extended experience*. *Neuron*, 2009. **63**(4): p. 497-507.
157. Gupta, A.S., et al., *Segmentation of spatial experience by hippocampal theta sequences*. *Nat Neurosci*, 2012. **15**(7): p. 1032-9.
158. Jackson, J.C., A. Johnson, and A.D. Redish, *Hippocampal sharp waves and reactivation during awake states depend on repeated sequential experience*. *J Neurosci*, 2006. **26**(48): p. 12415-26.
159. Masimore, B., J. Kakalios, and A.D. Redish, *Measuring fundamental frequencies in local field potentials*. *J Neurosci Methods*, 2004. **138**(1-2): p. 97-105.
160. Mitra, P. and H. Bokil, *Observed brain dynamics*. 2008, Oxford ; New York: Oxford University Press. xxii, 381 p.
161. Pesaran, B., et al., *Temporal structure in neuronal activity during working memory in macaque parietal cortex*. *Nat Neurosci*, 2002. **5**(8): p. 805-11.
162. Barnett, L. and A.K. Seth, *The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference*. *J Neurosci Methods*, 2014. **223**: p. 50-68.
163. Bressler, S.L. and A.K. Seth, *Wiener-Granger causality: a well established methodology*. *Neuroimage*, 2011. **58**(2): p. 323-9.
164. Seth, A.K., A.B. Barrett, and L. Barnett, *Granger causality analysis in neuroscience and neuroimaging*. *J Neurosci*, 2015. **35**(8): p. 3293-7.

165. Geweke, J., *Measurement of Linear-Dependence and Feedback between Multiple Time-Series*. Journal of the American Statistical Association, 1982. **77**(378): p. 304-313.
166. Geweke, J.F., *Measures of Conditional Linear-Dependence and Feedback between Time-Series*. Journal of the American Statistical Association, 1984. **79**(388): p. 907-915.
167. Seth, A.K., *A MATLAB toolbox for Granger causal connectivity analysis*. J Neurosci Methods, 2010. **186**(2): p. 262-73.
168. Box, G.E.P. and G.M. Jenkins, *Time series analysis; forecasting and control*. Holden-Day series in time series analysis. 1970, San Francisco,: Holden-Day. xix, 553 p.
169. Merchant, H., et al., *Cognitive modulation of local and callosal neural interactions in decision making*. Front Neurosci, 2014. **8**: p. 245.
170. Leuthold, A.C., et al., *Time series analysis of magnetoencephalographic data during copying*. Exp Brain Res, 2005. **164**(4): p. 411-22.
171. Adhikari, A., et al., *Cross-correlation of instantaneous amplitudes of field potential oscillations: a straightforward method to estimate the directionality and lag between brain areas*. J Neurosci Methods, 2010. **191**(2): p. 191-200.
172. Parent, M.A., et al., *Identification of the hippocampal input to medial prefrontal cortex in vitro*. Cereb Cortex, 2010. **20**(2): p. 393-403.
173. Colgin, L.L., *Oscillations and hippocampal-prefrontal synchrony*. Curr Opin Neurobiol, 2011. **21**(3): p. 467-74.
174. Heien, M.L. and R.M. Wightman, *Phasic dopamine signaling during behavior, reward, and disease states*. CNS Neurol Disord Drug Targets, 2006. **5**(1): p. 99-108.
175. Logothetis, N.K. and B.A. Wandell, *Interpreting the BOLD signal*. Annu Rev Physiol, 2004. **66**: p. 735-69.
176. Pfeiffer, B.E. and D.J. Foster, *Hippocampal place-cell sequences depict future paths to remembered goals*. Nature, 2013. **497**(7447): p. 74-9.
177. Mazur, J.E., *An adjusting procedure for studying delayed reinforcement.*, in *Quantitative analyses of behavior: The effect of delay and of intervening events on reinforcement value*, J.A.R. Nevin, H., Editor. 1987, Erlbaum: Hillsdale, NJ. p. 55-73.
178. Torres, L.V., et al., *Transitional and steady-state choice behavior under an adjusting-delay schedule*. J Exp Anal Behav, 2011. **95**(1): p. 57-74.
179. Cardinal, R.N., et al., *Local analysis of behaviour in the adjusting-delay task for assessing choice of delayed reinforcement*. Neural Netw, 2002. **15**(4-6): p. 617-34.
180. Blumenthal, A., et al., *Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task*. Neurobiol Learn Mem, 2011. **95**(3): p. 376-84.
181. Schoenbaum, G. and H. Eichenbaum, *Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex*. J Neurophysiol, 1995. **74**(2): p. 733-50.
182. Hassani, O.K., H.C. Cromwell, and W. Schultz, *Influence of expectation of different rewards on behavior-related neuronal activity in the striatum*. J Neurophysiol, 2001. **85**(6): p. 2477-89.
183. Roesch, M.R. and C.R. Olson, *Neuronal activity in primate orbitofrontal cortex reflects the value of time*. J Neurophysiol, 2005. **94**(4): p. 2457-71.
184. Zald, D.H., *The rodent orbitofrontal cortex gets time and direction*. Neuron, 2006. **51**(4): p. 395-7.
185. Rieke, F., *Spikes : exploring the neural code*. Computational neuroscience. 1997, Cambridge, Mass.: MIT Press. xvi, 395 p.

186. van der Meer, M.A., et al., *Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task*. *Neuron*, 2010. **67**(1): p. 25-32.
187. Madden, G.J. and W.K. Bickel, *Impulsivity : the behavioral and neurological science of discounting*. 1st ed. 2010, Washington, DC: American Psychological Association. xvi, 453 p.
188. Walton, M.E., et al., *Giving credit where credit is due: orbitofrontal cortex and valuation in an uncertain world*. *Ann N Y Acad Sci*, 2011. **1239**: p. 14-24.
189. Young, J.J. and M.L. Shapiro, *The orbitofrontal cortex and response selection*. *Ann N Y Acad Sci*, 2011. **1239**: p. 25-32.
190. Buzsáki, G., *Rhythms of the brain*. 2006, Oxford ; New York: Oxford University Press. xiv, 448 p.
191. Buzsaki, G. and B.O. Watson, *Brain rhythms and neural syntax: implications for efficient coding of cognitive content and neuropsychiatric disease*. *Dialogues Clin Neurosci*, 2012. **14**(4): p. 345-67.
192. Fell, J., et al., *Human memory formation is accompanied by rhinal-hippocampal coupling and decoupling*. *Nat Neurosci*, 2001. **4**(12): p. 1259-64.
193. Gregoriou, G.G., et al., *Long-range neural coupling through synchronization with attention*. *Prog Brain Res*, 2009. **176**: p. 35-45.
194. Hyman, J.M., et al., *Medial prefrontal cortex cells show dynamic modulation with the hippocampal theta rhythm dependent on behavior*. *Hippocampus*, 2005. **15**(6): p. 739-49.
195. Jones, M.W. and M.A. Wilson, *Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task*. *PLoS Biol*, 2005. **3**(12): p. e402.
196. Adhikari, A., M.A. Topiwala, and J.A. Gordon, *Synchronized activity between the ventral hippocampus and the medial prefrontal cortex during anxiety*. *Neuron*, 2010. **65**(2): p. 257-69.
197. Fries, P., *A mechanism for cognitive dynamics: neuronal communication through neuronal coherence*. *Trends Cogn Sci*, 2005. **9**(10): p. 474-80.
198. Harris, K.D., et al., *Organization of cell assemblies in the hippocampus*. *Nature*, 2003. **424**(6948): p. 552-6.
199. Mann, E.O. and O. Paulsen, *Role of GABAergic inhibition in hippocampal network oscillations*. *Trends Neurosci*, 2007. **30**(7): p. 343-9.
200. Cassenaer, S. and G. Laurent, *Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts*. *Nature*, 2007. **448**(7154): p. 709-13.
201. Benchenane, K., et al., *Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning*. *Neuron*, 2010. **66**(6): p. 921-36.
202. Markram, H., et al., *Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs*. *Science*, 1997. **275**(5297): p. 213-5.
203. Bi, G.Q. and M.M. Poo, *Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type*. *J Neurosci*, 1998. **18**(24): p. 10464-72.
204. Womelsdorf, T., et al., *Modulation of neuronal interactions through neuronal synchronization*. *Science*, 2007. **316**(5831): p. 1609-12.
205. Jutras, M.J., P. Fries, and E.A. Buffalo, *Gamma-band synchronization in the macaque hippocampus and memory formation*. *J Neurosci*, 2009. **29**(40): p. 12521-31.
206. Gruber, A.J., R.J. Hussain, and P. O'Donnell, *The nucleus accumbens: a switchboard for goal-directed behaviors*. *PLoS One*, 2009. **4**(4): p. e5062.

207. Colgin, L.L., et al., *Frequency of gamma oscillations routes flow of information in the hippocampus*. Nature, 2009. **462**(7271): p. 353-7.
208. Buzsaki, G., *Theta oscillations in the hippocampus*. Neuron, 2002. **33**(3): p. 325-40.
209. McCracken, C.B. and A.A. Grace, *Persistent cocaine-induced reversal learning deficits are associated with altered limbic cortico-striatal local field potential synchronization*. J Neurosci, 2013. **33**(44): p. 17469-82.
210. Courtemanche, R., N. Fujii, and A.M. Graybiel, *Synchronous, focally modulated beta-band oscillations characterize local field potential activity in the striatum of awake behaving monkeys*. J Neurosci, 2003. **23**(37): p. 11741-52.
211. Leventhal, D.K., et al., *Basal ganglia beta oscillations accompany cue utilization*. Neuron, 2012. **73**(3): p. 523-36.
212. Howe, M.W., et al., *Habit learning is associated with major shifts in frequencies of oscillatory activity and synchronized spike firing in striatum*. Proc Natl Acad Sci U S A, 2011. **108**(40): p. 16801-6.
213. Fries, P., et al., *Modulation of oscillatory neuronal synchronization by selective visual attention*. Science, 2001. **291**(5508): p. 1560-3.
214. Fries, P., et al., *The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4*. J Neurosci, 2008. **28**(18): p. 4823-35.
215. Donnelly, N.A., et al., *Oscillatory activity in the medial prefrontal cortex and nucleus accumbens correlates with impulsivity and reward outcome*. PLoS One, 2014. **9**(10): p. e111300.
216. Friston, K., R. Moran, and A.K. Seth, *Analysing connectivity with Granger causality and dynamic causal modelling*. Curr Opin Neurobiol, 2013. **23**(2): p. 172-8.
217. Kaminski, M., et al., *Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance*. Biol Cybern, 2001. **85**(2): p. 145-57.
218. Granger, C.W.J., *Investigating Causal Relations by Econometric Models and Cross-spectral Methods*. Econometrica, 1969. **37**(3): p. 424-438.
219. Salazar, R.F., et al., *Content-specific fronto-parietal synchronization during visual working memory*. Science, 2012. **338**(6110): p. 1097-100.
220. DeCoteau, W.E., et al., *Learning-related coordination of striatal and hippocampal theta rhythms during acquisition of a procedural maze task*. Proc Natl Acad Sci U S A, 2007. **104**(13): p. 5644-9.
221. van Wingerden, M., et al., *Theta-band phase locking of orbitofrontal neurons during reward expectancy*. J Neurosci, 2010. **30**(20): p. 7078-87.
222. Grothe, I., et al., *Switching neuronal inputs by differential modulations of gamma-band phase-coherence*. J Neurosci, 2012. **32**(46): p. 16172-80.
223. Gregoriou, G.G., et al., *High-frequency, long-range coupling between prefrontal and visual cortex during attention*. Science, 2009. **324**(5931): p. 1207-10.
224. Baldauf, D. and R. Desimone, *Neural mechanisms of object-based attention*. Science, 2014. **344**(6182): p. 424-7.
225. Nolte, G., et al., *Robustly estimating the flow direction of information in complex physical systems*. Phys Rev Lett, 2008. **100**(23): p. 234101.
226. Catanese, J., van der Meer, M., *Cortico-striatal and striato-cortico flow of information are associated with high and low gamma oscillations respectively*, in Annual Society for Neuroscience Conference. 2014, Society for Neuroscience: Washington, D.C.
227. Singer, W. and C.M. Gray, *Visual feature integration and the temporal correlation hypothesis*. Annu Rev Neurosci, 1995. **18**: p. 555-86.

228. Rajagovindan, R. and M. Ding, *Decomposing neural synchrony: toward an explanation for near-zero phase-lag in cortical oscillatory networks*. PLoS One, 2008. **3**(11): p. e3649.
229. Antzoulatos, E.G. and E.K. Miller, *Increases in functional connectivity between prefrontal cortex and striatum during category learning*. Neuron, 2014. **83**(1): p. 216-25.
230. Howe, M.W., et al., *Prolonged dopamine signalling in striatum signals proximity and value of distant rewards*. Nature, 2013. **500**(7464): p. 575-9.
231. Miller, E.K. and J.D. Cohen, *An integrative theory of prefrontal cortex function*. Annu Rev Neurosci, 2001. **24**: p. 167-202.
232. Guthrie, E.R., *Tolman on associative learning*. Psychological Review, 1937. **44**: p. 525-528.
233. Bett, D., et al., *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, 2012. **6**: p. 70.
234. Hu, D., X. Xu, and F. Gonzalez-Lima, *Vicarious trial-and-error behavior and hippocampal cytochrome oxidase activity during Y-maze discrimination learning in the rat*. Int J Neurosci, 2006. **116**(3): p. 265-80.
235. Griesbach, G.S., D. Hu, and A. Amsel, *Effects of MK-801 on vicarious trial-and-error and reversal of olfactory discrimination learning in weanling rats*. Behav Brain Res, 1998. **97**(1-2): p. 29-38.
236. Hu, D., G. Griesbach, and A. Amsel, *Development of vicarious trial-and-error behavior in odor discrimination learning in the rat: relation to hippocampal function?* Behav Brain Res, 1997. **86**(1): p. 67-70.
237. Hu, D. and A. Amsel, *A simple test of the vicarious trial-and-error hypothesis of hippocampal function*. Proc Natl Acad Sci U S A, 1995. **92**(12): p. 5506-9.
238. Schmidt, B., et al., *Conflict between place and response navigation strategies: effects on vicarious trial and error (VTE) behaviors*. Learn Mem, 2013. **20**(3): p. 130-8.
239. Muenzinger, K.F., *Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency*. Pedagogical Seminary and Journal of Genetic Psychology, 1938. **53**(1): p. 75-86.
240. Holland, P.C., *CS-US interval as a determinant of the form of Pavlovian appetitive conditioned responses*. J Exp Psychol Anim Behav Process, 1980. **6**(2): p. 155-74.
241. Breton, Y.A., K.D. Seeland, and A.D. Redish, *Aging impairs deliberation and behavioral flexibility in inter-temporal choice*. Front Aging Neurosci, 2015. **7**: p. 41.
242. Gold, J.I. and M.N. Shadlen, *The neural basis of decision making*. Annu Rev Neurosci, 2007. **30**: p. 535-74.
243. Tolman, E.C., *Prediction of vicarious trial and error by means of the schematic sowbug*. Psychological Review, 1939. **46**(4): p. 318-336.
244. Dember, W.N. and H. Fowler, *Spontaneous alternation behavior*. Psychol Bull, 1958. **55**(6): p. 412-28.
245. Sherrick, M.F., et al., *Rats' sensitivity to their direction of movement and spontaneous alternation behaviour*. Q J Exp Psychol, 1979. **31**(1): p. 83-93.
246. Graybiel, A.M., *The basal ganglia and chunking of action repertoires*. Neurobiol Learn Mem, 1998. **70**(1-2): p. 119-36.
247. Cromwell, H.C. and W. Schultz, *Effects of expectations for different reward magnitudes on neuronal activity in primate striatum*. J Neurophysiol, 2003. **89**(5): p. 2823-38.
248. Funahashi, S., M.V. Chafee, and P.S. Goldman-Rakic, *Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task*. Nature, 1993. **365**(6448): p. 753-6.
249. Niv, Y., et al., *Tonic dopamine: opportunity costs and the control of response vigor*. Psychopharmacology (Berl), 2007. **191**(3): p. 507-20.

250. Nicola, S.M., *The nucleus accumbens as part of a basal ganglia action selection circuit*. Psychopharmacology (Berl), 2007. **191**(3): p. 521-50.
251. Schweighofer, N., et al., *Low-serotonin levels increase delayed reward discounting in humans*. J Neurosci, 2008. **28**(17): p. 4528-32.
252. Miyazaki, K.W., et al., *Optogenetic activation of dorsal raphe serotonin neurons enhances patience for future rewards*. Curr Biol, 2014. **24**(17): p. 2033-40.
253. Yates, J.R., et al., *Role of medial prefrontal and orbitofrontal monoamine transporters and receptors in performance in an adjusting delay discounting procedure*. Brain Res, 2014. **1574**: p. 26-36.
254. Donaldson, Z.R., et al., *Genetic approaches for understanding the role of serotonin receptors in mood and behavior*. Curr Opin Neurobiol, 2013. **23**(3): p. 399-406.
255. Friedman, M., *Essays in positive economics*. 1964, Chicago ; London: University of Chicago Press. 328 p.
256. Green, L. and J. Myerson, *A discounting framework for choice with delayed and probabilistic rewards*. Psychol Bull, 2004. **130**(5): p. 769-92.
257. Ainslie, G., *Breakdown of will*. 2001, Cambridge ; New York: Cambridge University Press. xi, 258 p.
258. Kennerley, S.W., T.E. Behrens, and J.D. Wallis, *Double dissociation of value computations in orbitofrontal and anterior cingulate neurons*. Nat Neurosci, 2011. **14**(12): p. 1581-9.
259. Kim, S., et al., *Prefrontal and striatal activity related to values of objects and locations*. Front Neurosci, 2012. **6**: p. 108.
260. Stalnaker, T.A., N.K. Cooch, and G. Schoenbaum, *What the orbitofrontal cortex does not do*. Nat Neurosci, 2015. **18**(5): p. 620-7.
261. O'Doherty, J.P., *The problem with value*. Neurosci Biobehav Rev, 2014. **43**: p. 259-68.
262. Cisek, P., *Making decisions through a distributed consensus*. Curr Opin Neurobiol, 2012. **22**(6): p. 927-36.
263. Clayton, N.S., T.J. Bussey, and A. Dickinson, *Can animals recall the past and plan for the future?* Nat Rev Neurosci, 2003. **4**(8): p. 685-91.
264. Wassum, K.M., et al., *Phasic mesolimbic dopamine release tracks reward seeking during expression of Pavlovian-to-instrumental transfer*. Biol Psychiatry, 2013. **73**(8): p. 747-55.
265. Shiflett, M.W. and B.W. Balleine, *At the limbic-motor interface: disconnection of basolateral amygdala from nucleus accumbens core and shell reveals dissociable components of incentive motivation*. Eur J Neurosci, 2010. **32**(10): p. 1735-43.
266. Schoenbaum, G., et al., *Orbitofrontal Cortex and Outcome Expectancies: Optimizing Behavior and Sensory Perception*. Neurobiology of Sensation and Reward, ed. J.A. Gottfried. 2011, Boca Raton FL: Llc.
267. Schoenbaum, G., et al., *Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala*. Neuron, 2003. **39**(5): p. 855-67.
268. Takahashi, Y.K., et al., *Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning*. Neuron, 2013. **80**(2): p. 507-18.
269. Lucantonio, F., et al., *Orbitofrontal activation restores insight lost after cocaine use*. Nat Neurosci, 2014. **17**(8): p. 1092-9.
270. Lucantonio, F., D. Caprioli, and G. Schoenbaum, *Transition from 'model-based' to 'model-free' behavioral control in addiction: Involvement of the orbitofrontal cortex and dorsolateral striatum*. Neuropharmacology, 2014. **76 Pt B**: p. 407-15.
271. West, E.A., et al., *Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques*. J Neurosci, 2011. **31**(42): p. 15128-35.

272. Takahashi, Y.K., et al., *Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex*. Nat Neurosci, 2011. **14**(12): p. 1590-7.
273. Takahashi, Y.K., et al., *The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes*. Neuron, 2009. **62**(2): p. 269-80.
274. Cai, X. and C. Padoa-Schioppa, *Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex*. J Neurosci, 2012. **32**(11): p. 3791-808.
275. Cai, X. and C. Padoa-Schioppa, *Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation*. Neuron, 2014. **81**(5): p. 1140-51.
276. Stott, J.J. and A.D. Redish, *Representations of Value in the Brain: An Embarrassment of Riches?* PLoS Biol, 2015. **13**(6): p. e1002174.
277. Hunt, L.T., et al., *Mechanisms underlying cortical activity during value-guided choice*. Nat Neurosci, 2012. **15**(3): p. 470-6, s1-3.
278. Hare, T.A., et al., *Transformation of stimulus value signals into motor commands during simple choice*. Proc Natl Acad Sci U S A, 2011. **108**(44): p. 18120-5.
279. Redgrave, P., T.J. Prescott, and K. Gurney, *The basal ganglia: a vertebrate solution to the selection problem?* Neuroscience, 1999. **89**(4): p. 1009-23.
280. Czubyko, U. and D. Plenz, *Fast synaptic transmission between striatal spiny projection neurons*. Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15764-9.
281. Tunstall, M.J., et al., *Inhibitory interactions between spiny projection neurons in the rat striatum*. J Neurophysiol, 2002. **88**(3): p. 1263-9.
282. Frank, M.J. and E.D. Claus, *Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal*. Psychol Rev, 2006. **113**(2): p. 300-26.
283. Ahmari, S.E., et al., *Repeated cortico-striatal stimulation generates persistent OCD-like behavior*. Science, 2013. **340**(6137): p. 1234-9.
284. Burguiere, E., et al., *Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors*. Science, 2013. **340**(6137): p. 1243-6.
285. Kay, L.M. and W.J. Freeman, *Bidirectional processing in the olfactory-limbic axis during olfactory behavior*. Behav Neurosci, 1998. **112**(3): p. 541-53.
286. van der Meer, M.A. and A.D. Redish, *Theta phase precession in rat ventral striatum links place and reward information*. J Neurosci, 2011. **31**(8): p. 2843-54.
287. Bosman, C.A., C.S. Lansink, and C.M. Pennartz, *Functions of gamma-band synchronization in cognition: from single circuits to functional diversity across cortical and subcortical systems*. Eur J Neurosci, 2014. **39**(11): p. 1982-99.
288. Fries, P., *Neuronal gamma-band synchronization as a fundamental process in cortical computation*. Annu Rev Neurosci, 2009. **32**: p. 209-24.
289. Buzsaki, G., C.A. Anastassiou, and C. Koch, *The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes*. Nat Rev Neurosci, 2012. **13**(6): p. 407-20.
290. Sirota, A., et al., *Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm*. Neuron, 2008. **60**(4): p. 683-97.
291. Berke, J., *Participation of Striatal Neurons in Large-Scale Oscillatory Networks*, in *The basal ganglia VIII: Advances in behavioral biology*, J.P. Bolam, C.A. Ingham, and P.J. Magill, Editors. 2006, Springer Science & Business Media. p. 1-24.
292. Bracci, E., et al., *Voltage-dependent membrane potential oscillations of rat striatal fast-spiking interneurons*. J Physiol, 2003. **549**(Pt 1): p. 121-30.

293. Taverna, S., B. Canciani, and C.M. Pennartz, *Membrane properties and synaptic connectivity of fast-spiking interneurons in rat ventral striatum*. Brain Res, 2007. **1152**: p. 49-56.
294. Sharott, A., et al., *Different subtypes of striatal neurons are selectively modulated by cortical oscillations*. J Neurosci, 2009. **29**(14): p. 4571-85.
295. Cardin, J.A., et al., *Driving fast-spiking cells induces gamma rhythm and controls sensory responses*. Nature, 2009. **459**(7247): p. 663-7.
296. Sohal, V.S., et al., *Parvalbumin neurons and gamma rhythms enhance cortical circuit performance*. Nature, 2009. **459**(7247): p. 698-702.
297. Buzsaki, G., L.W. Leung, and C.H. Vanderwolf, *Cellular bases of hippocampal EEG in the behaving rat*. Brain Res, 1983. **287**(2): p. 139-71.
298. Klausberger, T., et al., *Brain-state- and cell-type-specific firing of hippocampal interneurons in vivo*. Nature, 2003. **421**(6925): p. 844-8.
299. Kim, T., et al., *Cortically projecting basal forebrain parvalbumin neurons regulate cortical gamma band oscillations*. Proc Natl Acad Sci U S A, 2015. **112**(11): p. 3535-40.
300. Tamas, G., et al., *Proximally targeted GABAergic synapses and gap junctions synchronize cortical interneurons*. Nat Neurosci, 2000. **3**(4): p. 366-71.
301. Carmichael, J., van der Meer, M., *Spatiotemporal distribution of ventral striatal local field potentials revealed by high-density silicon probes*, in *Annual Society for Neuroscience Meeting*. 2014, Society for Neuroscience: Washington, D.C.
302. Britt, J.P., et al., *Synaptic and behavioral profile of multiple glutamatergic inputs to the nucleus accumbens*. Neuron, 2012. **76**(4): p. 790-803.
303. Goto, Y. and A.A. Grace, *Limbic and cortical information processing in the nucleus accumbens*. Trends Neurosci, 2008. **31**(11): p. 552-8.
304. Chang, J.Y., P.H. Janak, and D.J. Woodward, *Neuronal and behavioral correlations in the medial prefrontal cortex and nucleus accumbens during cocaine self-administration by rats*. Neuroscience, 2000. **99**(3): p. 433-43.
305. Swanson, L.W., *Cerebral hemisphere regulation of motivated behavior*. Brain Res, 2000. **886**(1-2): p. 113-164.