Effects of rewards and reward-predictive cues on gamma oscillations in the ventral striatum

by

Sushant Malhotra

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2014

Author's Declaration

I hereby declare that I, Sushant Malhotra, am the sole author of this thesis. I understand that my thesis may be made electronically available to the public.

Abstract

Decisions, such as choosing between different rewards, are known to be influenced by a number of variables such as value, uncertainty and delay associated with a rewarding outcome. Various structures in the brain are responsible for handling different aspects of reward related decision making. To understand how such decisions are made, we can attempt to reverse engineer the brain. This involves understanding how brain activity is related to the representation and processing of rewards and also to subsequent behavior in response to rewarding events. One of the central elements of the reward circuitry of the brain is the ventral striatum. It has traditionally been known as the limbic-motor interface and thought to act as a link between various structures in the brain that are responsible for processing reward and reward related behavior. To study the neural processes that underlie processing rewards, I recorded from the ventral striatum of rats as they performed a cue-reward task. The aim of my project was twofold: First, to examine how rats behave in response to changes in value and uncertainty associated with a particular rewarding outcome and second, to investigate how rewards and cues that predict rewards are represented in the neural activity of the ventral striatum. Rats (n=6) were trained on a cue-reward task, where cues indicated the mean or variance of associated outcome distributions. Behavioral responses to the reward predictive cues demonstrated that the rats learned the value and risk associated with subsequent reward outcomes. Ventral striatal gamma oscillations are known to align to rewards in a variety of reward motivated tasks. However, it is not clear if these oscillations are associated with anticipation of obtaining the reward or the reward itself. In previous studies, reward delivery has been correlated with the anticipation of reward. In the current work, a delay is used to distinguish between anticipation of reward and the reward delivery itself. This is achieved by making the rats nose poke for a fixed time interval before the arrival of reward. The analysis presented in this thesis reveals that ventral striatal gamma oscillations occur both during the anticipation and delivery of reward, opening up the possibility of formal tests. They also align to arrival of cues that predict rewarding outcomes. This suggests that gamma oscillations might be essential for modulating behavior in response to cues and rewards both before and after reward delivery. Ventral striatum is ideally situated to modulate behavior in response to rewarding events. Past studies show that ventral striatal neural activity is associated with reward and reward motivated actions. However, as suggested by the research presented in this thesis, it is not clear what specific aspects of the decision making process can be attributed to the ventral striatum once learning in complete. Studying the ventral striatum is important because its malfunctioning is implicated in brain disorders such as drug addiction.

Acknowledgements

Table of Contents

List of Figures

List of Abbreviations

Medium spiny neuron - MSN

Fast-spiking interneuron - FSI

Subiculum - Sub

Entorhinal Cortex - EC

Ventral striatum - vStr

Ventral tegmental area - VTA

Temporal difference reinforcement learning - TDRL

Long term potentiation - LTP

Local field oscillation - LFP

Hippocampus - HC

Conditioned stimulus - CS

Unconditioned stimulus - US

Pavlovian-instrumental transfer - PIT

Chapter 1: Decisions, behavior and ventral striatum

1.1 Introduction

Making decisions is an integral part of who we are. In today's world, any given person makes several decisions in a single day. These might range from simple decisions such as what route to take to school to more difficult ones like what subject to major in. The decisions we make are made not in isolation but are a product of our interaction with our surrounding environment and our internal motivations.

Decision making is a complicated process that can be influenced by a variety of factors. Some of the factors include past experience, context, rewards or punishments, risk, effort and even delay associated with various choices. Many decisions we make in our daily lives are shaped by how we interpret and process various kinds of rewards. This provides motivation to study behavior with respect to rewards.

The concept of reinforcement (rewarding and aversive events) is central to the major theories of learning and behavior. Rewards include those salient objects in the environment (or actions) which are considered to be naturally rewarding or cues which predict naturally rewarding outcomes. Some of the naturally rewarding outcomes include food, drugs and procreation. Secondary rewards include various simple and complex cues that are associated with the above rewards such as money. Reward motivated decision making involves analyzing a number of relevant decision variables associated with different rewarding outcomes. Moreover, it also involves understanding the relationships that exists between primary and secondary rewards which can be fairly complex in real world scenarios.

Behavioral neuroscience aims to understand decision making in terms of the neural activity of the brain. Understanding the neural mechanisms constituting different behaviors is essential for devising drugs and therapies for diseases such as addiction and obsessive compulsive disorder in which our decision making machinery goes awry. Moreover, it is also important to examine how rewards and cues might bias us towards making irrational choices. Various experiments have been conducted to analyze the effects of cues on decisions. In one such experiment, a group of people were given a mug containing a warm or a cold drink and asked to rate how friendly a particular person was. Controlling for various factors, it was found that people holding the warm mug are more likely to find a person friendly or 'warmer' (Williams and Bargh, 2008). This obviously makes no rational sense from an economic standpoint but illustrates how various cues in our environment might affect our choices and decisions we make.

Many studies show that various aspects of rewards and cues that predict them are represented in the brain (Kable and Glimcher, 2009; Schultz, 2000). The research described in this thesis deals with elucidating the neural mechanisms that are involved in the processing of cues and rewards. Specifically, how cues and rewards are encoded in the neural oscillations of the ventral striatum (vStr) of rats. The vStr is an important part of the reward circuitry of the brain. It is strategically located to use contextual, sensory and affective information about the external environment provided by its input structures, to direct appropriate actions in response to salient events such as rewards or even aversive outcomes (Mogenson et al., 1980).

The vStr is involved in learning optimal actions in response to reward and cues associated with rewards (Humphries & Prescott, 2010). A number of studies have implicated vStr in the processing of various characteristics of reward such as value, risk, effort required to obtain reward and even the delay to the reward (Schultz et al., 1992; Roesch et al., 2009; Taha and Fields, 2005; Setlow et al., 2003; Sugam et al., 2013). The current work analyzes how vStr neural oscillations respond to rewards and cues. Thus, it involves training rats on a cue-reward association task and simultaneously recording from their vStr. The paradigms used to train the rats are presented below along with classical theories of learning with respect to cues and rewards. The role of vStr in processing and learning with respect to rewards is also examined in the sections below to provide motivation for the current research. Along with analyzing vStr neural activity the current work also investigates the effect of value and risk on the behavior of rats in a cue-reward task.

Understanding decision making is at the heart of a variety of disciplines such as economics, operations research, computer science, psychology and neuroscience. My research falls under the broader field of behavioral neuroscience which tries to understand the neural mechanisms that underlie various decision making processes. Thus it tries to investigate the decisions that people make as a function of the activity of neurons; which are the fundamental building block of the brain. The above offers an opportunity to understand decision making with a much higher resolution than afforded by abstract models in economics and operations research which focus on the actions/decisions of a particular individual and not the underlying brain processes.

1.2 Experimental behavior paradigms

The effect of rewards and cues on behavior has been studied extensively in the field of psychology and behavioral neuroscience. A number of experimental paradigms such as Pavlovian and instrumental conditioning are used to study behavior in the lab. The research described in this thesis involves conditioning rats to learn relationships between cues and rewards. The above falls under appetitive Pavlovian conditioning, which is explained below.

In appetitive Pavlovian conditioning, a stimulus or a cue such as a particular sound is paired with the presentation of a reward such as food or water. The delivery of reward is independent of the animal's response. The sound cue is by itself uninteresting to the animal but acquires salience when repeatedly paired with food. Pavlovian learning involves associating a cue with an outcome. This learning is generally accompanied by a conditioned response to the cue such as salivating or approach. The cue used for conditioning is called the conditioned stimulus (CS) while the reward is called the unconditioned stimulus (US).

In instrumental conditioning the animal has to perform an action (such as a lever press) either in response to the cue (CS) (or without a cue) to obtain food (US). Instrumental conditioning has been extensively studied in behavioral experiments involving a variety of animals especially rats (Dickinson et al, 1998). It can involve associating a stimulus with a response or an action with a particular outcome. Thus, unlike Pavlovian conditioning the delivery of reward is dependent on the animal's response. Even simple behavioral tasks generally consist of both Pavlovian and instrumental components.

An interesting crossover between Pavlovian and instrumental paradigms is the Pavlovian-instrumental transfer (PIT) effect (Holland, 2004). A rat is conditioned to receive a reward following a stimulus such as a light cue. The rat then is also trained on an instrumental task where for example it has to press a lever to obtain reward. After the rat has reached a certain level of performance in both conditioning tasks its tested on a hybrid of the above two tasks under extinction (Extinction refers to the fact that the rat doesn't get any reward in the new hybrid task although it did get rewards in the Pavlovian and instrumental tasks). In the hybrid task the rat has to lever press for food just like the instrumental task but it is also shown the stimulus from the Pavlovian task from time to time. Although there is no real relationship between the Pavlovian and instrumental tasks, showing the light stimulus increases the frequency with which the rat presses the lever in the hybrid task. Thus interestingly, the Pavlovian cue (light cue) changes the instrumental response of the rat even though the light cue should have no significance in the hybrid task.

Pavlovian-instrumental transfer is considered to affect a wide range of human behaviors such as drug addiction especially cue dependent relapse where the presence of the drug cues might trigger continued seeking of drugs. (Talmi et al., 2008; Ludwig, 1974). PIT is also an elegant demonstration of how various Pavlovian cues in our environment might prime us in subtle ways. This can even occur in contexts unrelated to the cues demonstrating that our actions might be directly affected by the cues present around us. Moreover, PIT can be specific and general. In specific PIT only the instrumental response which has the same reward or outcome associated with the Pavlovian cue is affected. Whereas, in general PIT, the instrumental response to multiple outcomes is modulated by the presence of the Pavlovian cue (Corbit and Balleine, 2005). Interestingly, PIT is not just a physiological response mediated by the presentation of the cue but rather is dependent on the information provided by the cue as a net predictor of reward. In an interesting study, Delamater (1995) designed an experiment where

noise (cue) predicted reward (sucrose solution). Once the rats had learnt the above Pavlovian relationship, the sucrose solution was presented in the absence of the cue (noise) also, thus reducing the predictive power of the cue. Most importantly, the less predictive cue did not show PIT effect after devaluation. Thus the above shows that the difference in behavior is a result of the predictive relationship of the cue to reward.

Thus the above examples of behavioral conditioning illustrate that an animal can be conditioned to learn a particular response to a previously neutral stimulus by pairing the stimulus with a positive or negative reinforcer such as food and shock respectively. I utilize a Pavlovian conditioning task to train my rats to associate different cues with both value and risk associated with particular reward outcomes.

1.3 Behavior and learning

As discussed in the previous section, there is plenty of experimental evidence for the effect of cues and rewards on behavior. Cues or rewards presented by themselves in an experiment might be static unchanging entities whereas the response of a subject to the same cues is shaped by a dynamic learning process. This learning is a well-studied aspect of various experimental tasks. Since the current work involves conditioning rats to associate cues and rewards it is helpful to examine if their learning can be understood using existing theories of learning and motivation. Two such theories are described below along with their possible explanation of reward motivated learning and behavior.

Learning, in a paradigm such as Pavlovian conditioning is described as the associations made between the stimulus and the reinforcer which lead to predictions or expectations. Animals detect when events occur unexpectedly, corresponding to errors in predictions of reward or other reinforcers. Thus, current theories of associative learning contend that just temporal contiguity between a stimulus and a reinforcer is not enough for learning, rather an error between the predicted reward and the actual reward is also required to ensure learning (Tobler et al., 2005). This error, which is generally known as reward prediction error forms the basis of many of the current learning theories in psychology which explain animal behavior in a variety of experiments involving both Pavlovian and instrumental conditioning.

An important theory used to describe association formation (learning) of conditioned (CS) and unconditioned stimuli (US) is the Rescorla-Wagner model (Rescorla and Wagner, 1972). The model describes a number of studies in which the animals use a cue to predict a particular reward outcome. According to the Rescorla-Wagner model the learning of association between a stimulus (CS) and a reward (US) depends on the extent to which the reward is predictable. Thus, the error in reward prediction drives the associative strength between the stimulus and the reward.

$$\Delta V = \alpha\beta(\lambda - V) \tag{1.1}$$

In equation 1.1, the left hand side represents the change in the associative strength between the CS and the US in a particular trial while the right hand side is the difference between maximum conditioning a reward can produce and the current associative strength between the reward and stimulus, multiplied by learning rate parameters alpha and beta which denote the learning rate and the intrinsic associability of the stimulus respectively. The learning rate parameters determine the speed of learning. 'λ' represents the limit of learning while 'V' represents the current associative strength between the CS and US on any given trial. Learning is complete when λ is equal to V.

The model predicts that an unpredicted reward will cause a much bigger change in the associative strength of the stimulus compared to a predicted one. The model is based on two principal assumptions that learning occurs only when rewards are unpredicted and the predictions due to different stimuli can be summed together to give the total prediction for a particular trial. The model is able to explain a large amount behavioral data such as blocking and conditioned inhibition. The former occurs when a stimulus is conditioned to predict a reward and then another stimulus is added to the experiment. It is seen that the animal forms no association between the new stimulus and the reward as the reward still occurs as predicted. Therefore there is no reward prediction error hence no learning. Despite being widely applicable the model suffers from a major limitation. It divides the decision making task into trials rather than some unit of time. Thus it is not able to explain decision making as a temporally continuous process.

As mentioned above, the Rescorla-Wagner model does not account for learning as a temporal process where conditioned and unconditioned stimuli are processed in time. One class of models which builds on the Rescorla- Wagner model and describes decision making as a temporal process was developed by Sutton and Barto (Barto, 1990) and is known as temporal difference reinforcement learning (TDRL). In TDRL model, the interaction between an agent and its environment is described by a set of states. Formally, an agent's goal is to maximize the expected future reward given a particular state. TD learning helps an agent learn an optimal value and policy function. The value function for a particular state refers to the expected return starting from that state while following a certain policy. A policy involves the agent mapping states to actions to maximize the expected reward. TD estimates the value function by taking the difference between the estimated value and the actual return. The update to the value function utilizes the difference of successive estimates of the value function. The simplest TD algorithm estimates the value (expected return) of a certain state as next reward plus the value of next state. The learning rule in this case is:

$$\delta_{t+1} = r_{t+1} + \gamma V_t(S_{t+1}) - V_t(S_t) \tag{1.2}$$

In equation 1.2, the left side represents the temporal difference error used to update the value function. 't' represents a sequence of states, V_t represents the predicted value of a given state at a certain time step 't', 'r' is the reward associated with a certain state 'S', 'gamma' is a discounting factor less than 1 as distant rewards are less important. The TDRL model continuously updates the state of the world (represented by a number of different possible states) at every time instant. The associative strength between the stimuli and reward is used to not only predict a future reward but also to make predictions due to the stimuli that will be available in the next time step with gamma discounting the future predictions.

The Rescorla-Wagner model and TDRL model are useful in explaining the behavior or learning exhibited by animals in tasks involving associations between cues and rewards. The major assumption of Rescorla Wagner model is that the strength of associative learning is proportional to the extent the reward or US is surprising. Thus, it postulates that an animal will gradually learn to associate a cue with a rewarding outcome given enough trials. Also, if the sequence or timing of cues is unpredictable then learning can occur even if the animal has already learned a particular task. Pavlovian value of rewards in a particular cue-reward task can explain the strength of the conditioned response such as the speed of approach to a particular reward. This difference in conditioned response can be used as an indicator of learning in a task where a number of different rewards are presented to an animal. Animals can also learn an arbitrary set of actions like pressing levers or running towards a certain location (policy) to obtain rewards. TDRL models make a prediction about the optimal strategy that can be used by an animal given a certain set of circumstances. This is based on estimating a value and optimal policy function. Thus, in various scenarios TDRL predictions can be compared with actual behavior and neural activity in the brain to explain how the learning process actually takes place. For example, TDRL models have been used to explain the activity of dopamine neurons in the ventral tegmental area that projects to the vStr. The activity of dopamine neurons reflects the computation of a temporal difference error signal which is thought to drive learning in various cue-reward experiments (Schultz et al., 1997).

The models discussed above only take into account the value of an unconditioned stimulus such as a reward. However value is a subjective quantity affected by various other variables that impact a decision. We know from everyday experience that we take into account the value of a certain outcome but also the risk associated with it. The delay (Hariri et al., 2006) or the effort (Phillips et al., 2007) associated with obtaining a reward also affects the perceived value of the reward. The subjective value of certain rewards/punishments is modulated by the context in which we encounter them. Thus a variety of different behavioral variables such as risk, time, effort are necessary to explain the behavior of a subject in a particular decision making task. This makes it necessary to study behavior with respect to various different decision variables and simultaneously create models that can help explain the learning mechanisms that might be used by a particular animal in a particular decision making situation.

1.4 Ventral striatum and behavior

This section describes the vStr and examines its role in learning, behavior and processing of cues and rewards, thus providing motivation for recording vStr oscillations in the current work. The vStr is an important node in the reward circuitry of the brain. It is a part of the basal ganglia, a group of interconnected nuclei that form topographically organized loops with cortex and thalamus and therefore are implicated in a variety of functions such as exploration, learning and behavior (Pennartz et al., 2009). The striatum can be divided further into three main parts: dorsolateral striatum, dorsomedial striatum and vStr. The borders between the above structures are not exact. The classification is based upon type of inputs received by each structure along with location. The striatum is predominantly composed of medium spiny neurons (MSN), GABAergic neurons which form 90-97% of entire striatum neuron population in a rat (Humphries and Prescott, 2010). The rest consists of different kinds of interneurons, among which the fast spiking interneurons (FSI) are the most numerous. The MSN's form the output population of the striatum and are GABAergic in nature (Humphries and Prescott, 2010). The vStr is further divided into core and shell. The core region is similar to dorsal striatum in its morphology and the regions it projects to. The shell region on the other hand is unique as it has smaller MSN's and is the only part of the striatum that has direct projections to structures outside the basal ganglia.
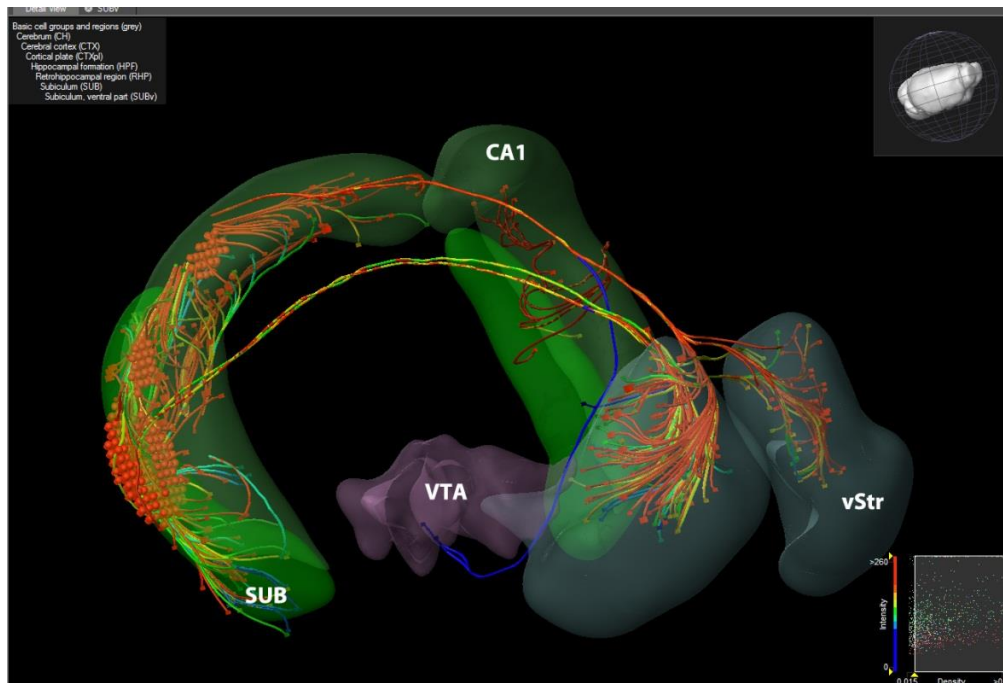

Figure 1.1: Anatomy of ventral striatum and its neighboring structures. The above figure (Website: ©2012 Allen Institute for Brain Science. Allen Mouse Brain Connectivity Atlas [Internet]. Available from: http://connectivity.brain-map.org)

shows some of the anatomical connections between the ventral striatum and its afferent structures for a mouse brain (very similar to a rat brain). In the figure different blocks of colors indicate different structures namely subiculum, ventral tegmental area and CA1 region of the hippocampus. The bright yellow and orange lines show neural projections from one area to another.

The vStr is a unique node in the reward circuitry of the brain as it is ideally situated to utilize contextual, affective, salient (pertaining to rewards) information about the environment to direct appropriate behavior (Cardinal et al., 2002). It receives a confluence of inputs from a variety of structures in the brain such as the prefrontal cortex, hippocampus, amygdala, and send outputs to motor control areas via thalamus (Humphries and Prescott, 2010). Structures such as hippocampus and amygdala are part of the limbic system, which is a set of primitive brain structures situated under the cortex, implicated in a variety of functions such as regulation of emotions, memory formation, reward motivation etc. Thus, vStr has been considered as a structure which acts as a limbic-motor interface (Mogenson et al., 1980). However the exact contribution of vStr to reward motivated behavior is not straightforward. This is partly because the vStr is a highly heterogeneous structure with multiple anatomical and functional dissociations including the division into core and shell, varying connectivity with structures outside the basal ganglia and even gradient in expression of receptors for different neuro-transmitters such as dopamine (Humphries and Prescott, 2010). The vStr is also innervated by dopamine neurons in the ventral tegmental area which modulate vStr activity in complex ways not easily mapped to behavior (Nicola, 2007).

Theories of learning and motivation emphasize the role of vStr in reward seeking behavior and in mediating the role of Pavlovian conditioned stimuli as well as unconditioned stimuli (Cardinal et al., 2002; Berridge and Robinson, 1998). Various studies have reported conflicting results with regards to the importance of vStr for responding to conditioned stimuli such as reward predicting cues. Some studies show that vStr lesions or reducing vStr dopamine profoundly impairs a rat's ability to respond to a cue in certain cue-reward tasks (Di Ciano et al. 2001; Parkinson et al. 2002). But many other tasks report that vStr lesions and dopamine manipulation have minimal effect on performance in instruction triggered tasks (Cole and Robbins, 1989; Giertler et al. 2004; Hauber et al. 2000). The above might be due to the differences in the times between two successive cue presentations (Nicola, 2007). Cues presented after long time intervals seem to require vStr while the same is not true for short time interval cues (Wadenberg et al., 1990, Robbins et al., 1990). Increasing dopamine in the vStr is sufficient to increase responding to reward predictive cues presented after long intervals (Nicola et al., 2005) and also increases conditioned responding during pavlovian-instrumental transfer (Wyvell and Berridge, 2000). Nicola and colleagues (2005) postulates that long interval cues are temporally more unpredictable from animal's point of view than short term ones and the former might require flexible approach strategy in order to obtain reward (Nicola, 2010).

It is important to note that most of vStr lesion and dopamine manipulation studies involve locomotion. Injection of dopamine agonists in vStr causes increased locomotion (Mogenson et al., 1980). Moreover vStr lesioned animals show deficits in approaching temporally unpredictable cues but not the predictable cues (Dalley et al. 2002; Parkinson et al. 2002). This suggests that vStr might be required for locomotion in certain kinds of cue responding. vStr dopamine is also related to effort expended to obtain rewards. Multiple studies by Salamone et al. (1999, 2005) show that high effort responding requires vStr dopamine while low effort responding doesn't. This might also be a reason for the cue responding deficits seen in certain tasks. Majority of vStr lesion and inactivation studies show that vStr seems to exert an inhibitory control on behavior by preventing an animal from committing a non-beneficial action (Cardinal et al., 2002). At the same time, certain studies show that vStr lesions result in reduced behavioral responding along with impaired responding to cues in certain cases (Balleine and Killcross, 1994; Corbit et al. 2001). Thus, vStr might be thought to modulate the impact of motivationally salient stimuli such as rewards and cues that predict them as well as invigorate actions to maximize reward gain.

Several recording studies have found reward and cue related neural activity in the vStr. Neurons in the vStr encode various aspects of rewards and cues that predict them (such as magnitude, palatability, probability) (McGinty et al., 2013, Taha et al., 2005, Setlow et al., 2003). For example, vStr neurons exhibit firing after the cue in instruction triggered tasks (Schultz et al., 1992). Such firing could be thought to invigorate a response to a reward predictive cue. Taha et al. (2006) found that a depression in activity of vStr neurons is necessary for reward consumption in rats and disturbing this activity via electrical stimulation prevents rats from sampling reward. Carelli and colleagues have conducted a number of experiments to study vStr neural activity in response to appetitive and aversive outcomes. Their studies provide evidence for separate vStr circuits encoding natural (water) and drug (cocaine) rewards (Carelli et al., 2000), encoding of costs associated with rewards by vStr neurons (Day et al., 2011) and encoding of rewarding and aversive taste stimuli by vStr neurons (Roitman et al., 2005). Ventral striatum dopamine concentration (dopamine neurons in the VTA project to vStr) has also been tied to various aspects of reward related behavior namely vStr dopamine release is related to effort and delay related reward costs (Day et al., 2010). Ventral striatal neural activity reflects discounting of rewards based on delay (Roesch et al., 2009). Roesch et al. have found that cue-responsive vStr activity is influenced by both expected reward and the response required to obtain the reward. Moreover, the neural activity is correlated with the speed at which the rats execute a response to obtain a reward indicating that vStr might be involved in integrating reward and motor information to aid decision making. Recently, McGinty and colleagues (2013) found that a subset of vStr neurons increased their activity temporally aligned with a reward predictive cue on a flexible approach reward-cue task. The same neurons were not active when the rat initiated movement to collect reward. However, their neural activity still predicted the onset of locomotion and the speed of approach to the reward location. The above result is in contrast to that reported by Goldstein et

al. (2012) who didn't see any correlation between vStr neural activity and the vigor of subsequent movement in an inflexible (rat commits the same action to obtain reward) approach cue-reward task. The above contrast in the results may be accounted by the flexible vs. inflexible approach required in the two tasks to obtain reward. This suggests that vStr modulation of behavior in response to cues is affected by a number of factors that might be attributed to the action required to attain reward or more generally the context in which the reward is obtained. Thus, vStr neural activity reflects various aspects of rewards and cues that predict rewards. Moreover, vStr neural activity is also correlated to the subsequent movement to obtain a reward.

Besides responses to cues and rewards, a salient characteristic of many MSN's found in several electrophysiological studies is a ramp-like increase in their firing rate in anticipation of a delivery of a reward (Shibata et al., 2001; Mizumori et al., 2004; van der Meer and Redish, 2009, Schultz 1992). These cells known as 'ramp' cells were first found while recording from the vStr of a monkey in tasks where reward delivery was predicted by specific cues (Apicella et al. 1991; Schultz et al. 1992; Hollerman et al. 1998). Recently these cells have received further scrutiny in reward tasks involving spatial navigation in rats (van der Meer and Redish, 2011). It is not clear what exactly ramp cells represent but they seem to respond to various different aspects of rewards such as value, identity (Miyazaki et al., 1998), ramping to cues predicting rewards, as well as distributed encoding of reward value with some cells increasing while others decreasing their firing rate in response to various reward outcomes (Khamassi et al., 2008). It is also been shown that these cells can ramp over time as well as space (Khamassi et al., 2008). Since these cells gradually increase their firing rate to various rewards, they are thought to represent a state value signal which fits in well with Temporal-difference reinforcement learning theories used to describe vStr neural activity (Khamassi et al., 2008). There is also some evidence from fMRI studies linking reduced ramping activity in the vStr with neurological disorders such as ADHD and depression (Epstein et al., 2006, Beck et al., 2009).

Local field potential oscillations (LFP) in the vStr are thought to be critical for gating the flow of salient information (occurrence of rewards) from limbic afferents (such as HC & amygdala) to basal ganglia outputs (Gruber et al., 2009; van der Meer et al., 2010). Beta and gamma oscillations are the most prominent oscillations throughout the striatum (van der Meer et al., 2010). Ventral striatal local field oscillations have been shown to exhibit distinct relationships to ongoing behavior especially actions in response to rewards (Howe et al., 2011; Berke et al., 2009; van der Meer and Redish, 2009; Kalenscher et al., 2010). Therefore, these rhythms are considered to be involved in processing of various aspects of rewards. Leventhal and colleagues (2012) found that presentation of an instruction cue on a Stop-Signal task was always followed by an increase in beta power (discrete burst of beta oscillations). However there was a second abrupt increase in beta power only if the rat correctly responded to the instruction cue. Moreover Howe and colleagues (2011) showed that for rats trained on a T-

maze task, gamma bursts were prominent during learning while beta bursts became more prominent once learning was complete. In addition, vStr MSN's and interneurons synchronized their firing to the peaks and troughs of the beta oscillations respectively (Howe et al., 2011).

Various vStr recording studies have reported gamma oscillations with two distinct bands namely gamma-50 and gamma-80 (Berke et al., 2009; van der Meer and Redish, 2009). Gamma-50 and gamma-80 tend to occur in bursts of 150-200 ms. Both, gamma-50 and gamma-80 show distinct relationships to reward motivated behavior, vStr inputs and single neuron spiking activity (van der Meer et al., 2010). Gamma-50 and gamma-80 show increase in power with respect to rewards. Gamma-50 has been shown to increase abruptly at reward sites while gamma-80 power ramps up gradually to reward sites (van der Meer and Redish, 2009). In contrast, Berke and colleagues (2009) found that gamma-50 power was abolished following reward receipt while gamma-80 showed a transient increase in power. Van der Meer and Redish (2009) also found that gamma-50 power increased post reward delivery. This was only true for rewarded trials and in line with the results of Kalenscher et al. (2010) who found that gamma-50 power was higher on rewarded trials compared to unrewarded ones. The above contrast in multi-phasic response of gamma oscillations to rewards might be attributed to difference in behavior on rewarded and unrewarded trials. However, the above studies do show that vStr gamma activity is modulated by rewards and affected by reward anticipation and receipt.

It is speculated that gamma oscillations might enable the vStr to act as a switchboard by allowing activity from certain input structures to affect vStr output and its subsequent effect on motivation and behavior (Gruber et al., 2009). Ventral striatum receives dense input from dopamine neurons in the VTA. Systemic injection of DA agonists such as amphetamine (which leads to excessive locomotion in rats) leads to switching from gamma-50 to gamma-80 (Berke et al., 2009). Moreover it has been shown that vStr neurons phase lock to gamma oscillations, indicating that these rhythms are related to vStr neural activity and are not just volume conducted from a nearby structure such as the piriform cortex (Berke et al., 2009; van der Meer and Redish, 2009; Kalenscher et al., 2010). Thus, gamma oscillations probably reflect processing of input activity by the vStr to guide behavior in response to rewards. The research presented in this thesis analyzes the relationship between vStr neural activity, specifically gamma oscillations, and processing of cues and rewards.

There is also evidence from human EEG and fMRI studies linking vStr activity to processing and anticipation of rewards. Cohen and colleagues (2009a) recorded local field potentials from the vStr of human subjects performing a reward task. They found the occurrence of alpha (4-12 Hz) waves and bursts of gamma (40-80 Hz) oscillations resembling those found in the rat vStr (van der Meer et al., 2010). Also the phase of the alpha waves at which the gamma bursts occurred differed between rewarded and unrewarded trials and strategy switches following losses were predicted by a breakdown of this cross frequency coupling (Cohen et al., 2009b). There have also been a number of fMRI studies showing that vStr activity is related to

anticipation of rewards (Knutson et al., 2001), action monitoring in response to rewards (Münte et al., 2007) and even related to salient behaviorally relevant events besides rewards (Zink et al., 2003). Thus there exists ample evidence from human vStr recordings showing the vStr is essential for the processing of rewards and also for initiating appropriate behavior in response to rewards.

Thus, in summary, vStr is involved in mediating the motivational effects of various CS and US along with processing the various characteristics associated with a particular US (food reward) such as identity, value, delay to the reward and even the effort costs associated with obtaining a reward. It is also necessary for obtaining rewards in a spatial environment using a flexible action strategy. Administration of certain drugs into vStr can affect behaviour in a variety of different ways such as inducing locomotion and conditioned reinforcement for a particular action.

1.5 Research project: description, motivation and goals

Past studies on the vStr have focused on the activity of single neurons in response to CS and US. The majority of the studies have focused on the effect of value and identity of different rewards on vStr neurons (Carelli, 2004; Taha and Fields, 2005; Setlow et al., 2003). As previously discussed, vStr activity is modulated by the value of rewards, expectation of reward and also the delay to particular rewarding outcomes (Schultz et al., 1992; Roesch et al., 2009) . Past studies also report that vStr neurons respond to cues which predict rewards (Setlow et al., 2003). Thus most of the recording literature on vStr has focused on value as a decision variable i.e. a variable that is utilized by the subject such as a rat to make reward related decisions.

However, value is not the only decision variable that affects decisions. Another decision variable of primary importance is uncertainty or the probability associated with different appetitive or aversive outcomes (Rangel et al., 2008; Platt and Huettel, 2008). When the exact probabilities associated with an uncertain outcome are known then it is considered a risky outcome. If the probabilities are unknown it is considered an uncertain outcome. fMRI studies of vStr (Sabrina et al., 2007; Kuhnen et al., 2005) have found that vStr activity is correlated with risk associated with a particular choice. However, there have been no vStr LFP recording studies done exploring the effect of risk as a decision variable. Risk here refers to the variance associated with a particular reward outcome. For example, if a cue predicts availability of either 1 pellet or 5 pellets each with 0.5 probability then the variance (risk) associated with the cue and the predicted reward is the variance of 1 and 5 which is 4. In my experiment (spatial reward-cue task), I vary both value and risk associated with a particular reward outcome while recording vStr local field potentials (aggregate neural activity in an area). The above provides insight into how vStr processes value and risk and how vStr LFP's represent salient aspects of

a reward outcome. Thus, the current work also helps to elucidate how vStr LFP's might represent aspects of reward predictive cues and the behavioural actions taken in response to the same cues. Furthermore, past LFP recording studies do not distinguish between anticipation and receipt of reward. The above two scenarios are separated in my experiment by including a fixed delay for which the rat needs to nose poke in order to receive reward.

Studying vStr neural responses to reward can open possibilities for better understanding of how behaviour is modulated by rewards and cues that predict them. More specifically, studying how different decision variables such as value and risk are encoded in the vStr neural activity might lead to more realistic neural implementations of various models of human decision making. Studying neural mechanisms underlying vStr activity can also help to elucidate how inputs from HC and PFC might affect reward processing in vStr. This work can form the basis of future studies which look at how ensembles of vStr neurons encode different aspects of a reward motivated decision making task. Analysis of vStr LFP's as well as simultaneous HC recordings can also inform future studies about how vStr inputs effect processing of rewards and subsequent actions in response to rewarding stimuli. All of the above can help us to decipher the neural processes that are responsible for various reward motivated decisions made by a particular organism. Finally, improper functioning of the vStr is an implicated in number of disorders such as addiction and obsessive compulsive disorder (Scheres et al., 2007; Beck et al., 2009). Better understanding of vStr function will help to develop more effective therapies for the aforementioned disorders.

Chapter 2: Neurophysiology

2.1 Neuron and neural code

To elucidate the neural mechanisms that underlie decision making with respect to cues and rewards, I recorded neural activity from the vStr of a rat performing a reward motivated task. The processing and interpretation of neural activity involves understanding how such activity is generated in the brain. In the following section, I describe the processes underlying the electrical brain signals recorded in my experiment and also techniques to analyze the recorded brain activity.

A neuron is the basic functioning block of the nervous system. It is an electrically excitable cell which communicates with other cells via electrical and chemical signaling (via specialized chemical connections called synapses). Each neuron has a cell body (soma), dendrites (inputs) and an axon (output). The dendrites are thin finger like projections from the cell body or soma. They are the major source of input to a particular neuron. The axon on the other hand carries nerve signals away from the soma via synapses where neurotransmitter chemicals (or electric signals) are released to communicate with the dendrites of other neurons. A voltage gradient is maintained across the neuronal membrane by a difference in the concentrations of various kinds of ions primarily sodium, potassium and chloride ions. Changes in concentrations of ions occur with the help of ion pumps and ion channels (which can be voltage gated) embedded in the membrane. The membrane of all neurons contains various kinds of ion channels, which can be open or closed. Many ions channels selectively allow certain kind of ions (for example: sodium, potassium, chloride, calcium) to pass through them. The purpose of ion channels is to allow the neuron to control the entry and exit of particular ions and prevent entry of others (semi-permeable). The membrane is said to be depolarized if the membrane voltage changes in the positive direction, typically due to the opening of ligand-gated ion channels after glutamate release by a presynaptic neuron. If the membrane gets depolarized to a certain threshold, the voltage can rise rapidly forming a positive feedback loop (due to the presence of voltage gated sodium ion channels) and then drop quickly. This electrochemical event is known as an action potential (Kandel et al., 2000). It can lead to activation of neighboring neurons via chemical release from the synapse of the neuron which is activated. This stereotyped change in membrane voltage is also referred to as a spike. Spikes are important because for majority of neurons, sub threshold changes in membrane potential do not get transmitted i.e. only the spikes are transmitted from one neuron to another via synapses. Thus generating spikes is the primary method by which neurons communicate with one another.

Neurons in the brain are collectively involved in the regulation of various processes such as recognition/processing of sensory inputs, detection and analysis of rewards (also other salient features in the environment), motor planning/movements or action selection which are

necessary for an organism's survival. In order to understand the functional principles that underlie the working of the brain, specifically processes involving cognition and behavior, it is helpful to understand the neuronal code i.e. the patterns of spikes using which neurons communicate with each other to perform complex analysis and computations. Various techniques exist to record the activity of neurons from different parts of the brain. Some of them concentrate on characterizing the behavior of a single neuron while others try to elucidate the functioning of the brain as a whole. Single neuron recording methods include techniques such as patch clamp recordings in which intracellular membrane voltage changes of a single neuron are studied in vitro or in vivo. Activity of several neurons can be studied in vivo by recording extracellular membrane currents from the neurons of a live behaving animal. This extracellular recording of neural activity can help to relate some variable in the external environment with the activity of the neurons in a particular part of the brain. Finally techniques such as functional magnetic resonance imaging (fMRI), Positron emission tomography (PET) enable researchers to study the activity of neurons in multiple areas of the human brain. This is vital to understand the brain on a systems level and to identify the parts of the brain that might be involved in a certain behavior or sensory process. All techniques involve tradeoffs related to temporal/spatial resolution and invasiveness. Extracellular recordings provide access to the activity of individual neurons unlike fMRI which is more suitable for studying the activity of the entire brain.

For the work in this thesis, extracellular recording was used to access neural processes involved in the processing of rewards and cues. This allowed me to relate behavior to the recorded neural activity. The above was achieved by implanting electrodes in the vStr of a behaving rat as it collected rewards. The recorded neural activity was filtered to obtain two kinds of signals. The first was the spikes or action potentials. This is the high frequency portion of the electrical signal. The low frequency portion of the electrical signal is known as the local field potential. This mainly represents the sum of dendritic synaptic activity of several neurons in the area around the electrode (several 100um) along with other slow changing signals such as hyperpolarization currents and volume conduction (Sirota et al., 2008). Both kinds of neural activity along with their advantages and limitations are further explained below.

2.2 Extracellular recording: description, advantages and limitations

As mentioned above, various recording techniques have their own advantages and limitations. Since the research described in this thesis involves recording extracellularly, it is important to note the underlying neural processes represented by such recordings and the information that can be deciphered from such recordings. The neural processes thought to underlie extracellular recordings are described below.

In extracellular recordings, potentials are measured across two points in the extracellular space rather than across the neuronal membrane. The extracellular space consists of a conductive fluid which allows the passage of current (ions) through it. When the neuron is at rest (uniformly polarized), there is no net current flow across its membrane. However an active neuron is non-uniformly polarized and might have its soma and dendrites at different potentials. This results in the flow of ions from the one part of the neuron to the other through the extracellular space. This passage of currents creates electric fields that exert force on any charge in their vicinity. This force can be measured as the extracellular potential. In an intracellular recording, the potential is measured across the membrane of the neuron itself by impaling it with a glass micro-electrode. Figure 2.1 below shows the waveform of a recorded intracellular action potential from a CA1 pyramidal neuron. The corresponding extracellular potential is shown as well.
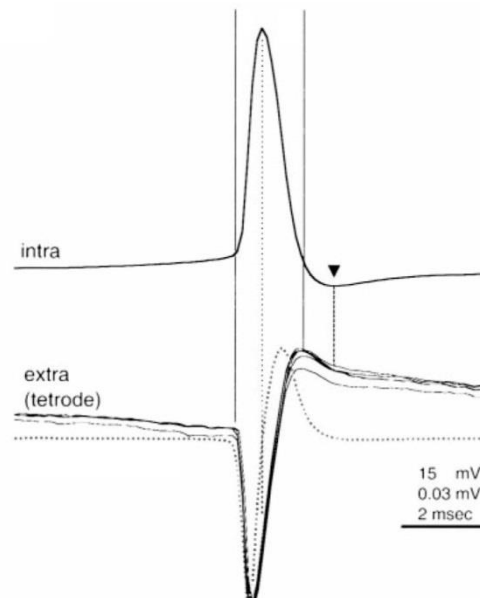
Figure 2.1: Waveform of an intracellular and extracellular action potential. The above figure adapted from Henze et al., 2000 shows both an averaged intracellular action potential (top) and the corresponding extracellular potential (bottom) recorded from a CA1 pyramidal cell in the hippocampus. The y-axis is voltage in mV while the x-axis is time. The intracellular recording was done by impinging the neuron with a glass pipette while the extracellular recording is done with a tetrode situated in the vicinity of the neuron. The figure shows how the intracellular potential of the neuron first increases from baseline as it depolarizes (opening of voltage gated sodium channels) and then subsequently decreases as the potassium channels open finally dropping below baseline before returning back to normal.

An active neuron can be considered as an electric dipole with the region where the action potential is initiated as the sink (negative charge) and the other parts as the source (positive charge). Thus, potential at any point in extracellular space will be due to electrical fields exerted at that point by various neurons (approximated as dipoles). Since extracellular potentials are measured between two points in the extracellular space the setup usually consists of a recording electrode (near the neurons of interest) and a distant reference electrode. Synaptic inputs to various locations on a neuron also set up electric dipoles. Since these inputs are constantly changing along with the activity of the neuron the extracellular signals recorded are not stationary in either time or space. Therefore, a recorded extracellular potential can reflect the activity of many neurons and can be very difficult to interpret since the electrical activity of many neurons affects the potential at a point in extracellular space, it is generally not easy to predict the potential at a certain point since the neurons are generally packed in complicated geometries. (Johnston and Wu, 1994)

If an action potential is initiated at a particular location in the axon of a neuron than the conductance change (conductance increases due to the opening of voltage gated ion channels especially sodium channels) at that point makes it a sink (produces an inward current) while another part of the neuron acts as the source. The time course of this trans-membrane current, dictates the time course of an extracellular potential recorded near a particular point. This is true for slow voltage changes which have the same time course as the membrane time constant. For fast changes, such as action potentials, the time course of the extracellular field potential is roughly proportional to the first derivative of the trans-membrane potential (Henze et al., 2000). Thus the extracellular potential is related to changes in membrane voltage (intracellular events) by the above two rough approximations. Thus by recording the extracellular voltage at a certain point one can roughly determine the fast spiking activity of the neurons in the area as well as the collective slow membrane voltage changes of many neurons in the brain tissue near the recording electrode. Extracellular recordings offer the opportunity to simultaneously record the activity of hundreds of neurons with fine temporal resolution. This is invaluable in understanding how networks of neurons perform behaviorally relevant computations by communicating with one another.

Once the extracellular voltages have been recorded, they go through a process called spike sorting which involves grouping spikes into clusters representing the activity of putative neurons. Spike sorting involves filtering the signal (high pass filter for spikes), detecting spikes (using some kind of threshold voltage), extracting distinguishable features of different spikes such as peak amplitude, width, energy, waveform shapes etc. and clustering the spikes with similar features into separate clusters. Spike sorting usually involves a subjective component where the experimenter has to validate or slightly alter the clusters created by the spike sorting algorithm. Thus it can be time consuming and not completely accurate. However by using a tetrode (four electrodes wound together) approximately 10-12 neurons in the vicinity can be detected as conflicting results from one channel can be resolved by using information from

other 3 channels (Henze et al., 2000). A basic way to do the above is to plot the peak amplitude of the four channels against one another and assess the amount of separation between different clusters. (Harris et al., 2000; Schmitzer-Torbert et al., 2005)

Local field potentials are the slow aggregate voltage fluctuations recorded near the electrode tip. They are thought to represent the presynaptic processes on the dendrites such as excitatory and inhibitory post synaptic potentials (Mitzdorf, 1985). Also soma-dendritic processes such as subthreshold membrane oscillations (Kamondi et al., 1998), afterpotentials and volume conduction are other processes that contribute to local field potentials. The exact spatial extent of the population of neurons contributing to LFP's varies from 200 µmm (Katzner et al., 2009) to few millimeters (Sirota et al., 2008). A limitation of LFP's is that their activity cannot be interpreted precisely as it reflects average changes in synaptic inputs and sub-threshold soma-dendritic voltage changes. LFP's also differ depending on the location of the electrode tip with respect to neurons and the geometry of the neurons (as discussed above since each neuron can be thought of as a dipole, different packing geometries lead to different LFP's). LFP oscillations have been observed in number of different brain structures at various frequencies such as theta (6-10 Hz), beta (15-25 Hz) and gamma (40-100 Hz). These oscillations are thought to be critical for functions such as coding of relevant information and inducing synaptic plasticity (Buzsáki, 2002), for motor preparation and planning (Sanes and Donoghue, 1993) and for processing of reward related information. Thus LFP could be thought of reflecting some sort of internal state of a network of neurons which are anatomically and functionally related to one another.

Another viewpoint of LFP's suggests there they are not just an epiphenomenon but are actually necessary for communication between various networks of neurons in the brain via coherence (Fries, 2005). This hypothesis is based on the fact that neurons when activated tend to exhibit rhythmical oscillations which are related to both spike output and the sensitivity to input for a given neuron population. It is postulated that only coherently oscillating groups of neurons can communicate with one another. For this to occur, the frequency, relative phase and conduction delay between two populations of neurons should be reliable. This would enable the oscillatory activity of one group to systematically excite or inhibit the activity of another group. Thus, coherence between different set of LFP's might provide a substrate for effective neural communication between different structures in the brain. The above however is a topic of active debate (Levine et al., 1999). It is postulated that anatomical connections and non-linear interactions between activities of several neurons might lead to local rhythmic patterns of electrical activity, recorded as local field oscillations. Thus, local field oscillations might arise as a byproduct of the structural and functional properties of neural sub-circuits in different parts of the brain and may not serve any particular purpose.

By studying vStr LFP's it is possible to access the activity several neurons. Analyzing this activity while the rats perform a certain reward motivated behavior provides an opportunity to study the neural mechanisms that might be important for such behavior. To record these analog

neural signals ones needs to implant electrodes into the brain which are then connected to a data acquisition device. This is explained further in the next chapter.

2.3 Analysis of recorded neural data

Ventral striatum exhibits distinct modes of processing with very different neural and LFP activity patterns. Examples of the above include theta, beta and gamma oscillations. These rhythms are thought to subserve a number of important functions related to behavior and synaptic plasticity underlying learning and behavior (van der Meer et al., 2010). They also have distinct relationships to intrinsic membrane properties and spiking activity of neurons found in the vStr (Berke et al., 2009, Kalenscher et al., 2010, van der Meer and Redish, 2009). Studying these distinct states of processing can provide valuable insight into the functioning of vStr. A number of different statistical techniques have been utilized to analyze the content of the aforementioned neural and LFP patterns. A list of techniques with an overview of their application to various vStr recording data from previous rat studies is presented below.

2.3.1 Overview of LFP data and analysis techniques

Power Spectrograms

Signals such as waves or oscillations can have energy associated with them. A signal's power is defined to be energy associated with it per unit time. Power spectral density refers to the strength of energy fluctuations in a given signal as a function of its component frequencies. Power spectral density can be used as a tool to study LFP's. This can be done by plotting power spectral density of an oscillation over time known as a power spectrogram. A power spectrogram can provide information about any changes in a particular frequency band of the LFP which might be related to a salient event in the external environment such as the appearance of cue or reward. It might also be correlated to behavior and input from other structures. Power spectrograms are one of the tools used to analyze the frequency content of various oscillations found in the vStr.

LFP recordings from the rat vStr reveal high power in two distinct frequency bands: one centered around 50 Hz referred to as gamma-50 and the other ranging from 70-90 Hz (Berke et al., 2009; van der Meer and Redish, 2009) referred to as gamma-80. Gamma oscillations also have specific behavioral correlates. Gamma-50 and gamma-80 show increase in power with respect to rewards. Gamma-50 has been shown to increase abruptly at reward locations while gamma-80 power ramps up gradually to reward sites (van der Meer and Redish, 2009).
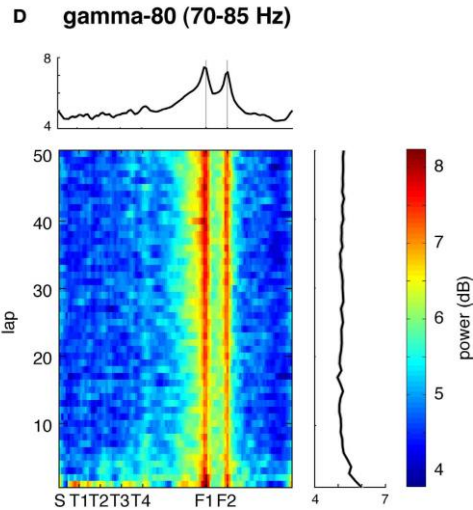
Figure 2.2: Spectrogram showing average power over gamma-80 frequency range (70-85 Hz) as the rat performs a spatial reward task. X-axis is position on the track (adapted from van der Meer and Redish, 2009). The rat starts at point S and is rewarded at feeders F1 and F2. The top panel shows average gamma-80 power over laps (rewarded laps only) during a certain session while the bottom panel shows lap by lap change in gamma-80 power (Y axis is lap number). It can be seen from the figure that gamma-80 ramps up slowly to reward sites and returns to baseline relatively quickly.

Coherence

Coherence is a measure of the similarity between the phase and the spectral power content of two time varying signals (such as local field oscillations recorded at two different locations) in a particular frequency band. A high coherence for two LFP's recorded in two different structures indicates that they are related in some manner, thus implying a functional or anatomical connection between the two structures.

Past studies have shown that there is a distinct relationship in the theta power range (8-10 Hz) between HC and vStr during reward approach (van der Meer and Redish, 2011). A plot of theta band coherence between the LFP's recorded in the two structures shows that coherence ramps up to reward sites as shown in figure 2.3 below.
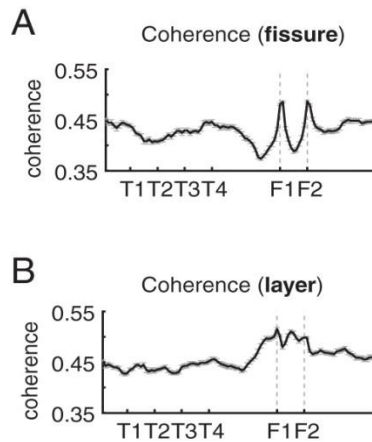
Figure 2.3: Theta Coherence between vStr and HC. The above figure (adapted from van der Meer and Redish (2011) shows the coherence in the theta band (6-10 Hz) between two different LFP pairs. Figure A shows the coherence between an LFP in the vStr and a LFP in the hippocampal fissure while figure B shows the coherence between a LFP in the vStr and a LFP in the Hippocampal layer. Theta coherence between vStr and HC tends to ramp up to reward sites F1 and F2.

Coherence is not used to analyze neural activity in the current work but it's discussed to convey its importance in assessing the plausibility of various hypotheses related to LFP's such as communication through neural coherence theory (Fries, 2005).

2.3.2 Overview of spike data and analysis techniques

This section provides an overview of the methods utilized in this thesis to analyze the spiking data recorded from the vStr. It also reviews recorded vStr neural data from previous studies.

Tuning Curves

Tuning curves are used to characterize the response of a neuron relative to some external variable. Generally they involve plotting the neuron's firing rate response on the y axis while the parameter is plotted on the x-axis.

While recording from the vStr of a rat, van der Meer and Redish (2011) found ramp neurons which gradually increased their firing rate to a reward site. Figure 2.4 is a plot of the spatial tuning curve of a vStr ramp neuron. The x-axis is space and the y axis is firing rate at a particular point in space.
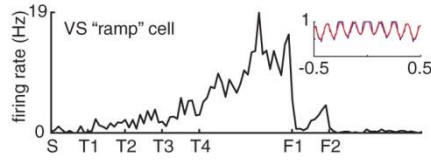
Figure 2.4: Spatial Tuning curve of a vStr neuron (adapted from van der Meer and Redish, 2011). In this figure, F1 and F2 are two different reward sites associated with the same amount of reward. The VS ramp neuron increases its firing rate as the rat approaches the first feeder F1. When the rat reaches feeder F1, the firing rate drops and increases again to feeder F2.

PETH (Peri-event time histogram)

A peri-event time histogram is a histogram of times a particular neuron fires with respect to some external event or stimulus. Thus PETH's are used to understand the relationship between timing and rate of neural spikes and some event over a number of repeated trials. The x-axis represents the time bins aligned to the onset of the event or stimulus while y axis shows the spike counts occurring in those time bins. (Brown et al., 2004)

Several studies have examined the relationship between the PETH's of vStr neurons and presentation of rewards or cues which predict those rewards (Roitman et al., 2005; Schultz et al., 1992). Carelli et al. have reported cue-responsive neurons which show changes in their firing rate time-locked to the presentation of reward predictive cues (Roitman et al., 2005). As discussed previously, a subset of vStr neurons show a depression in their activity during reward intake (Taha et al., 2006). Altering this neural activity by using stimulation prevents the rat from consuming the reward, thus implying some sort of causal relationship between vStr neurons and reward motivated behavior (Taha et al., 2007).

The below figure shows a PETH of an anticipatory ramp neuron in the vStr where y axis represents the spike counts while the x axis represents time bins. Time '0' is the pellet release time for a particular feeder in the task. It can be seen in the figure that the neuron ramps up to the arrival of pellets as the rat approaches the feeder.
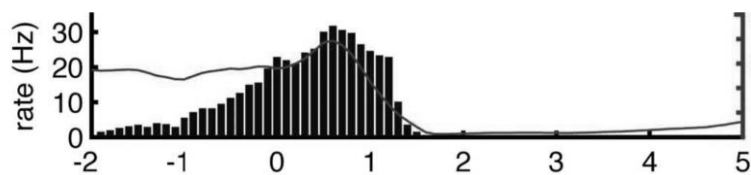

Figure 2.5: PETH of a vStr ramp neuron (adapted from van der Meer and Redish, 2009a).

Thus there is ample evidence that vStr neurons encode several relevant characteristics of rewards and cues. These include value, identity of reward, delay to reward, effort expended to obtain reward (Schultz et al., 1992; Setlow et al., 2003; Roitman et al., 2005). However these past studies have generally focused on single neuron activity and not analyzed the relationship between LFP's and rewards and cues that predict rewards, which is the main objective of this thesis.

LFP-spike phase relationships

Several different techniques are used to determine the relationship between the LFP's and the spikes recorded in a certain part of the brain. The aim of analyzing LFP-spike relationships is twofold, firstly to understand if the LFP has any local relevance (i.e. related to activity of neurons found nearby); secondly to see if LFP-spike relationships play any role in neural coding of relevant stimuli and rewards. LFP-spike phase locking involves calculating the particular phase of the LFP at which a neuron fires and representing the same in phase histograms. LFP-spike triggered average is another method where the acquired LFP trace is averaged centered on the spike times of a particular neuron. If there is no relationship between LFP and the spikes of the neuron the average comes out to be zero, else it shows some systematic variability with respect to the spike times.

These methods can provide insight into the functional relationship between single neural activity and LFP's. Past studies have shown that vStr gamma and neural activity are intimately related. Activity of single neurons (both MSN's and FSI's) is related to both gamma power and phase (Berke et al., 2009, Kalenscher et al., 2010, van der Meer and Redish, 2009). Many FSI's cohere with gamma-50 or gamma-80 at different points during a certain task (Berke et al., 2009). Thus the above results indicate that part of vStr gamma is locally generated and may be relevant to processing of reward information. A plot of a phase histogram of a vStr MSN neuron with respect to gamma rhythm is shown in figure 2.6 below.
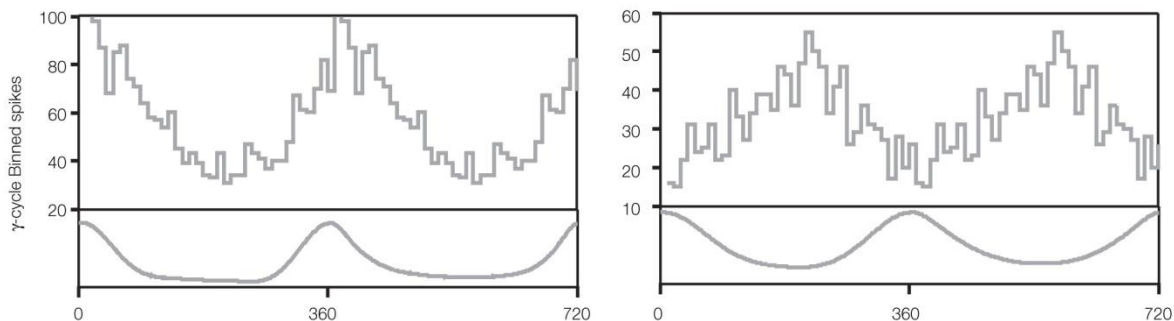
Figure 2.6: Phase histograms of two vStr MSN's (reprinted from Kalenscher et al., 2010).

The phase histograms of two putative medium spiny neurons are shown in figure 2.10 above (reprinted from Kalenscher et al., 2010). The graphs show the activity of neurons as a function of the phase of the gamma LFP cycle (30-100 Hz). The top panel shows the average number of spikes per gamma phase bin (bin size: 10 degrees) while the bottom panel display the gamma cycle. The first neuron fires at a preferred phase of 0 degrees while the second one fires at a preferred phase of about 180 degrees. Both neurons show clear phase locking to gamma LFP.

Chapter 3: Experimental design and methods

This chapter describes the setup, equipment and the techniques utilized to obtain the neural data analyzed in this thesis. The primary aim of this thesis is to examine the ventral striatal LFP (local field oscillations) activity underlying the processing of value and risk associated with cues and rewards. For this purpose, rats were trained to perform a cue-reward task. Electrodes were subsequently implanted in the vStr of the trained rats to record the activity of vStr neurons and local field potentials as the rats performed a reward motivated task. Various statistical techniques were used to analyze the relationship of the recorded LFP neural activity to cues, rewards and behavior.

3.1 Subjects

The experiments were carried out using 6 Long Evans rats aged between 4-10 months at the start of training. These rats were trained on two separate reward tasks described later in the section. The rats were food deprived to no less than 85% of their free eating body weight so as to motivate them to seek out food. They were also housed in a colony room with a 12 hour light/12 hour dark cycle.

3.2 Setup

Rats were trained to run back and forth on a linear track shown in figure 3.1 below. The track was constructed out of wood and painted black. Reward receptacles (basins where rewards were deposited for rats to eat) were made out of plastic. These receptacles were connected to Coulbourn H14-23R automated pellet dispensers that released TestDiet 5TUL food pellets on either end of the track. Each pellet weighed 0.045 grams. A camera mounted on the ceiling directly above the track was used to keep track of the location of the rat during the task. The experimental setup also included two speakers, one on each end of the track. The speakers were used to play the cues (audio tones) while the rat approached the corresponding feeder. Coulbourn H20-94 infra-red sensors were installed at feeder receptacles on either end to record the time the rat actually entered his nose into the receptacle to collect the reward. The experimental setup was controlled using a MATLAB script which interfaced with the Neuralynx data acquisition system (Digital Lynx) equipped with digital input/output ports. Neuralynx's input/output ports were used to continuously monitor the activity of the camera and the infra-red sensors to release the pellets at appropriate times as the rat ran back and forth

on the track. The above is further described in the next section. The track along with the feeders is shown in figure 3.1 below.

Figure 3.1: Linear track including the feeder receptacles.

Figure 3.1 above shows the linear track along with the two feeders on either side. The track was 2.5 meters in length and 12 centimeters off the ground. The camera used to track the rat while it runs on the track is directly above the track (not shown on in the figure). The infra-red nose poke sensors are attached to the black feeder receptacles. The feeders are connected to the receptacles using a hollow plastic tube. The speakers are attached to the two feeder stands on either end. A feeder and the corresponding feeder receptacle are indicated by a yellow and red arrow respectively.

3.3 Task and training

Rats were trained to run back and forth on the track to collect pellets on either end. The track was divided into three different zones as seen by the software. When the rats entered the middle zone (the location of which was jittered to ensure that rats couldn't predict exactly when a cue would be played) an auditory cue was played that predicted the value of the reward the rat would receive or the risk associated with the future reward. Additionally the rats had to nose poke for a specified time period of 0.5 seconds on either end in order to receive the reward. This delay helped to control for running speed confounds and helped in differentiating neural signals pertaining motion from those that were related to rewards. Figure 3.2 shows the series of events leading up to the dispensing of pellets from the pellet dispenser.
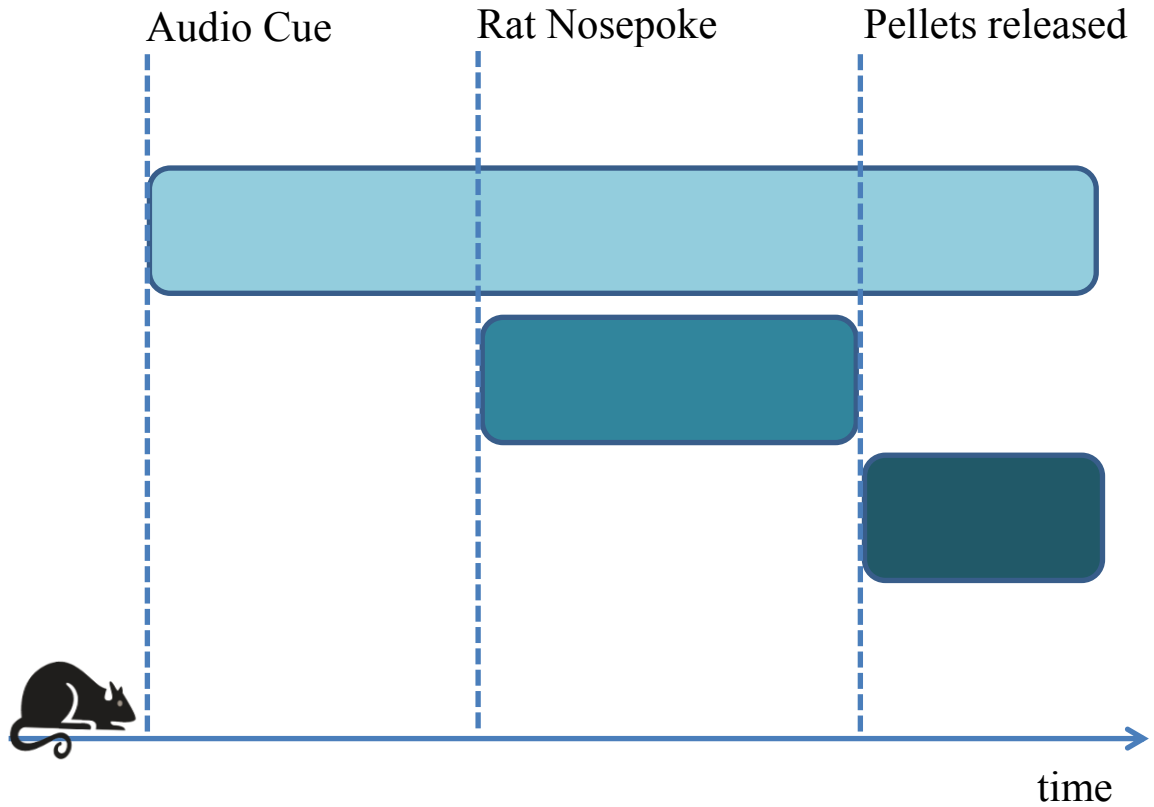
Figure 3.2: This block diagram shows the progression of example trial as the rat approaches a particular pellet dispenser on either end of the track. The audio cue is played at a random location near the middle of the track. The rat then has to nose poke at the pellet receptacle for 0.5 seconds (once completely trained) to trigger the dispenser to release pellets. This process is repeated at every trial with rat alternating between either ends of the track. The audio cue stops playing once the last pellet is dispensed or 10 seconds after its onset, depending on which of the two scenarios occurs first.

There were five different audio cues used: low frequency tone, high frequency tone, white noise, siren1 and siren2. The intensity of the played cues was balanced across the two speakers to ensure the same loudness (dB) level of approximately 75 dB at the track. This was done by testing the right and left speakers cue decibels levels using a sound level meter. A description of the cues is given below:

S1- Low frequency tone - 2 kHz tone, turning on/off at 10Hz

S2- High frequency tone - pure 15 kHz tone

S3- White noise

S4- Siren 1 - 8 kHz tone, sine wave amplitude-modulated at 2Hz

S5- Siren 2 - 3 different mixed sinusoidal tones alternating at 15Hz

Each of the above cues was randomly assigned to one of the five different reward outcomes (explained further below) for a particular rat. The outcome-cue associations for all rats are given in the figure 3.3 below:

Rat	Low reward	Medium reward	High reward	Low risk	No risk	High risk
R002	S1	S2	S3	-	-	-
R011	S3	S2	S4	S1	S2	S5
R014	S4	S1	S5	S3	S1	S2
R016	S3	S2	S4	S1	S2	S5
R018	S2	S5	S1	S4	S5	S3
R020	S5	S4	S2	S3	S4	S1

Figure 3.3: Audio cues corresponding to reward outcomes associated with different values and risks for all rats. The top row indicates the type of reward outcome while the leftmost column lists all the rats trained on the task.

The cues were counterbalanced across rats. Thus, for a given rat the association between cue and outcome was fixed, but across rats this relationship was randomized. There were five different reward outcomes namely 1, 2, 3, 4 or 5 pellets. A particular session of the task on a single day was divided into blocks: Value and Risk, each of which was 20 minutes in duration. The order of the blocks was generated randomly. Rat 2 was the only rat who didn't have the risk block and ran only on the value task for 40 minutes instead.

In the value section, the cues predicted the value of the upcoming reward which was limited to 1, 3 or 5 pellets. Thus cues low reward, medium reward and high reward (which were chosen randomly for a given rat) predicted 1, 3 and 5 pellets respectively. The order of cue presentation was randomized within blocks of 30 cues each. The low and high reward cues were presented with a probability of 0.4 each while the medium reward cue was presented with a probability of 0.2. The above assignment of probabilities ensured that all cues were presented approximately the same number of times (the medium reward cue was presented again in the risk task but not low and high reward cues). The rats had to run back and forth (alternate between the two ends of the track) to receive pellets Each rat started training with a nose poke delay of 50 ms. The rats were considered nose poke trained once they could reliably nose poke for 500 ms. If the rat missed a particular nose poke (exited before the experimenter specified nose poke delay) no reward was given and the rat would have to start a new trial by running to the opposite side of the track.

In the risk block, the cues predicted the risk associated with the upcoming reward which was again limited to three outcomes: low risk, no risk and high risk. In the current task, Risk was defined to be the variance associated with the predicted reward outcomes. This definition of risk has been used in previous studies (O'Neill and Schultz, 2010). The 3 pellet cue (medium

reward) from the value section was used again to represent no risk (variance of 0) while two new cues were used to represent low and high risk. The 3 pellet cue indicated a reward of 3 pellets every time and thus the associated reward outcome had no variance. The reward outcomes associated with the low risk cue were 2, 3 or 4 pellets (variance of 0.8). The 2 and 4 pellet reward outcomes could occur with a probability of 0.4 each while the 3 pellet reward outcome occurred with a probability of 0.2. The reward outcomes associated with the high risk cue were 1, 3 or 5 pellets (Variance of 3.2). Just like the low risk cue, the 1 and 5 pellet reward outcomes could occur with a probability of 0.4 each while the 3 pellet reward outcome occurred with a probability of 0.2. The particular probabilities associated with the different reward outcomes ensured that the average reward for all three cues was the same (3 pellets) but the variance was different as described above. Also the low risk cue and high risk cue were presented with a probability of 0.4 each while the no risk cue was presented with a probability of 0.2. A pictorial illustration of the cues along with their associated reward outcomes is shown in the figure 3.4 below. The numbers indicate the probability of a given reward outcome.
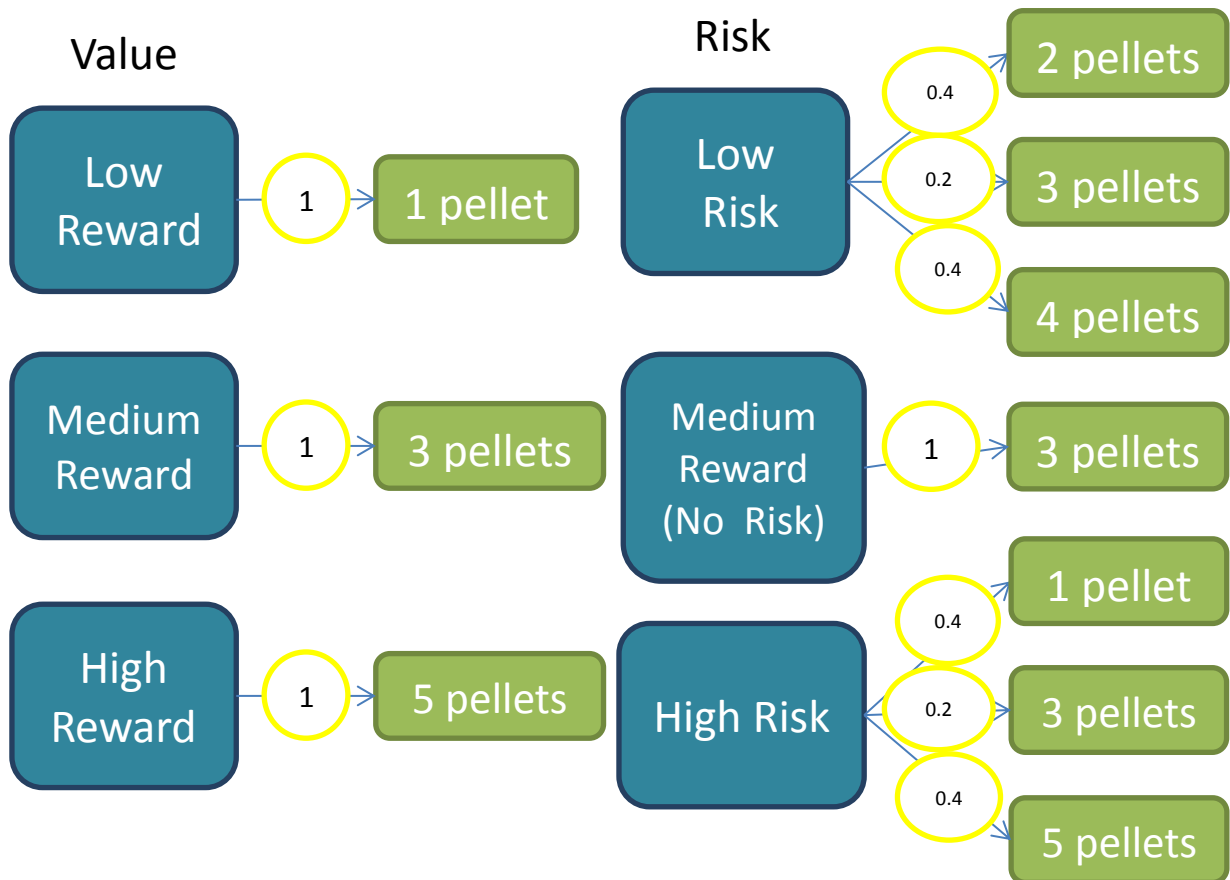

Figure 3.4: This figure depicts the relationship between the cues (blue blocks) for both the value and risk blocks and the number of pellets associated (green blocks) with each cue, including the probabilities (yellow circles) associated with particular reward

outcomes. In the value block, cues low reward, medium reward and high reward predicted rewards of 1, 3 and 5 pellets respectively each time. In the risk block, cue medium reward still predicted 3 pellets every time. But cue low risk predicted 2 and 4 pellet reward outcomes with a probability of 0.4 each and the 3 pellet reward outcome with a probability of 0.2 (low risk). Similarly cue high risk predicted 1 and 5 pellet reward outcomes with a probability of 0.4 each and the 3 pellet reward outcome with a probability of 0.2 (high risk).

The task timeline was also divided into three phases. The handling phase, training phase and the recording phase. During handling the rats were made familiar with the human experimenter. This involved the rats being handled by the experimenter for 10 minutes daily for at least 10 days. After that the rats entered the training phase, during which they were put on the track. The main aim of this phase was to allow the rats to learn the value and risk tasks. Once the rats had achieved reasonable performance (50 laps in both sessions) along with the ability to distinguish between the cues they were considered trained. The above was confirmed by plotting the rats running speed centered on the onset of 3 cues in the value task. This running speed indicated the speed of approach to the feeders to collect rewards. If the rats running speed with respect to the different cues differed significantly and if they ran faster for more pellets for at least 3 consecutive training sessions, they were considered to have learned the relationship between the cues and the subsequent rewards. At this point the rats were ready for surgery and were implanted with recording electrodes. After recovery, the rats were put back on the track to record vStr neural signals while they performed both the value and risk task. The above steps are outlined in figure 3.6.

3.4 Surgery

After the rats had achieved performance during training, they were implanted with a dual-bundle hyperdrive (an array of independently movable tetrodes) directed towards CA1 (HC) and vStr. (CA1 targets: 3.8 mm posterior and 2.5 mm right-lateral from bregma, two tetrodes; vStr targets: ~1.6 mm anterior and 2.0 mm right-lateral to bregma, four tetrodes, one reference). Tetrodes used in the implant were constructed from 90/10% Pt/Ir (17 micron diameter, insulated with polyamide) California Fine Wire. They were plated with platinum-black to impedances in the 250-300 kohm range.
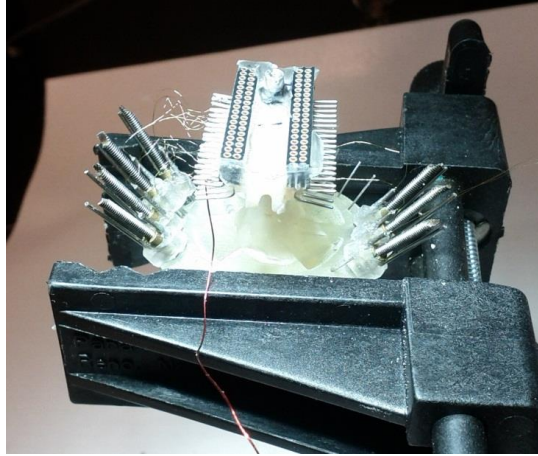
Figure 3.5: A hyperdrive during construction.

Figure 3.5 shows a hyperdrive under construction with tetrodes coming out of cannulas attached to shuttles which can translate up or down in the vertical direction. The hyperdrive was chronically affixed to the rat's skull during surgery and remained attached over the duration of the experiment.

The surgery was performed by anesthetizing rats with sodium pentobarbital (45mg/kg, i.p. following induction with isoflourane). During surgery, anesthesia was maintained using isoflurane (0.5-2% in oxygen). The rats were also injected with Anafen (analgesic) and dualcillin (antibiotic) to prevent any infections. The breathing and the temperature of the rat was monitored for any abnormalities throughout the surgery (every 15 minutes). A minimum of two people performed each surgery with one being the sterile surgeon and the other assisting the former. All the instruments used in the surgery were autoclaved to prevent any infections. Rats were secured in a stereotactic apparatus throughout the duration of the surgery. The HC and vStr targets were located on the skull relative to bregma. A standard 1.8mm trephine was used to make both the HC and vStr craniotomies. Holes were drilled into the skull with a standard 0.7 mm burr. Jewelers screws (FST, 00-90, 1/8" long) were inserted into the drilled holes to affix the implant onto the skull. Dental cement was used to cement the hyperdrive to the screws. Rats were given a post-surgery injection of Baytril (antibiotic) with saline to prevent infections and to keep them hydrated. Also, their condition was monitored until they were awake. All procedures were approved by the animal care committee at the University of Waterloo (AUPP #11-06), and were in accordance with CCAC guidelines for animal care. The rats were allowed to recover for 4-7 days following surgery. They were given food ad libitum and ibuprofen in their drinking water. Rats were also given Baytril injections under anesthesia for two days following surgery. They were anesthetized by using isoflurane (2% in oxygen). Once the rat was unconscious, Baytril was injected subcutaneously. Care was taken to minimize suffering or any subsequent infections.

3.5 Recording

Neural activity was recorded using a Digital Lynx recording system (neuralynx). This system functioned as an Analog to Digital Converter and simultaneously recorded spike as well as LFP activity. Only those spikes that met a threshold value set by the experimenter were recorded. Putative neurons were isolated by using MClust 3.5 (AD Redish, University of Minnesota) after automatic pre-clustering using KlustaKwik 1.5 (KD Harris, Imperial College London). To reduce noise, the spike/LFP data was referenced against the ventral striatal reference. A baseline recording of 5 minutes duration was done just before and after the experiment to assess the stability of the recording on a particular day.

The hyperdrive was connected to an electronics board known as the headstage or preamplifier board. This was used to amplify the electrical signals obtained from the electrodes implanted inside the brain. The headstage was then connected to a Data Acquisition System (DAQ), a multi-channel Analog to Digital Converter (ADC). The ADC sampled the electrical activity received from the electrodes at 32 kHz for spikes (2 kHz for LFP's) and converted it into a digital format that could be read by the acquisition software and finally imported into MATLAB for statistical analysis. The DAQ also controlled and received data from various other parts of the experimental setup such as the camera and was used to monitor the rat's position, feeders and sensors.

3.6 Perfusion and histology

After completion of recording, the location of each tetrode was marked using a 10 second, 0.2 mA current passed through the tetrode to produce a microlesion (gliosis). These marks were used aid future histological investigation to determine the exact position of the tetrodes in the brain. The animals were anesthetized using isoflurane and subsequently euthanized using a CO_2 chamber. Perfusion was performed immediately following euthanization. The rats were perfused with formalin and the brains were removed from the skull and stored in a formalin-sucrose solution. Coronal sections of the brain were cut using a cryostat and they were stained with cresyl violet The tetrode tracks were investigated using a microscope. Only those tetrodes which were in vStr and HC were used for further analysis of LFP's or isolation of neurons.

3.7 Flowchart of project timeline

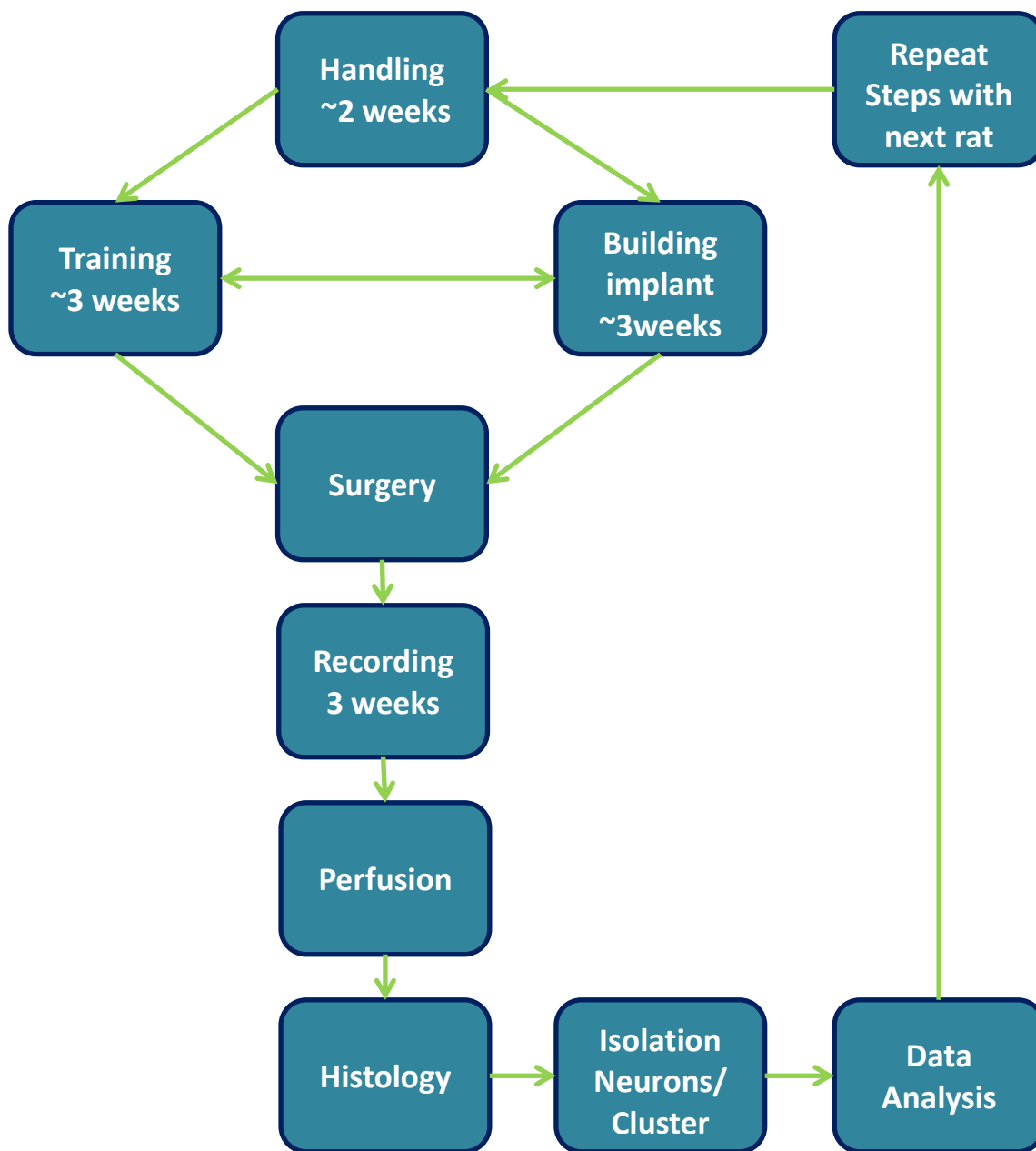The flowchart below shows the timeline of the experiment for a particular rat.

Figure 3.6: Flowchart showing the timeline of the experiment for a particular rat. Shows the timeline of various phases involved in training and recording from a rat.

3.8 MATLAB techniques

To analyze the neural and LFP activity in response to cues and reward outcomes I used a number of different statistical techniques implemented in MATLAB. Some of these techniques are used only to analyze spikes while others to analyze LFP's. The application of some of these techniques to past vStr neural data has already been discussed in Section 2.

Peri-event time histogram

To analyze if vStr neurons respond to cues and different reward outcomes in a systematic manner, I made peri-event time histograms centered on cue and reward delivery times. I used a custom PETH script implemented in MATLAB. This function divided the experimental period (time) into bins of specified size and then created a histogram of the spikes in those bins around the event of interest.

Power spectral density estimation

I also investigated LFP activity at times of presentation of cues and rewards. This was to understand if decision variables such as value and risk were encoded by changes in LFP spectral power content.

Electrophysiological recordings are prone to noise from various sources such as line noise or charge build up. This can show as slow fluctuations in the recorded LFP voltage signal. To remove this noise, I used the locdetrend function of the Chronux toolbox. This function utilizes a 1 s moving window to select particular time samples of the signal. The best fitting line (using least squares) for each sample is weighted and then used to obtain the estimate of the signal noise which is then removed (Mitra and Bokil, 2007). Also, the LFP data was interpolated to a fixed time scale so as to prevent any intermittent changes in sampling rate from affecting the results.

To actually calculate average power spectral density for each 20 min session, a spectrogram was constructed for individual vStr LFP's. This was done using a multi-taper method which calculated the spectral density in a finite sliding time window by computing the discrete Fourier transform of the given time series (in our case the sampled LFP data) (Mitra and Pesaran, 1999). The function used for this purpose was the mtspecgramc of the Chronux 2.0 toolbox (http://**chronux**.org/) with the following parameters: window size, 0.5 s; time step, 50 ms, five tapers. This function estimated the spectral density by using the below formula:

$$S(f) = 1/K \sum_{k=1}^{K} |x_k(f)|^2$$

$$x_k(f) = \sum_{t=1}^{T} w_t(k)x_t e^{-2ift}$$

where S is the averaged spectral density, $w_t(k)$ are the tapers, x_t is the signal

Tapers were used to calculate the discrete fourier transform since the LFP data was finite. However, a discrete fourier transform requires an infinite time series. Taper or a windowing function smoothly decreases the values of the time series being analyzed to zero on either end of the tapered interval. By moving this taper across the sample time series a more reliable estimate of spectral power can be calculated by reducing broadband bias (mixing of far away frequencies).

Chapter 4: Results

4.1 Behavior

Six rats R002, R011, R014, R016, R018 and R020 were trained on the linear track cue task described in the Methods chapter. Briefly, this task paired audio cues with different reward outcomes. The value and risk associated with a particular audio cue was varied in a systematic order. Once the rats learned the task they were implanted with recording electrodes in their hippocampus and the vStr. The task consisted of a value and risk block, in which rats were cued (using an audio tone) about the value and the risk associated with the upcoming reward outcome respectively. All rats completed both blocks except R002, who was only run on the value block. Rats were required to nose poke at a feeder receptacle for a specified delay (described below) in order to receive the reward. This enabled us to analyze vStr neural activity in response to expectation of reward without confounds due to running speed, posture and other kinds of movement. The nose poke requirement also helped to dissociate any motion related artifacts from reward related activity. The association between the cues and the reward outcomes were counterbalanced across rats so any biases associated with a particular cue were averaged out across animals.

It is well known that rats run faster when approaching large rewards compared to small rewards (Crespi, 1942). Thus, to determine if the rats learned the relationship between the tone cues and the reward outcomes, I analyzed the running speeds of the rats in response to the presentation of different cues. All rats initially showed random responses to various cues but eventually ran fastest for the five pellet cue and slowest for the 1 pellet cue in the value block as shown in the figures below.

4.2 Rats discriminate between cues predicting different reward values

The figures below show the behavior of rats during both value and risk task. The plots (PETH) show the running speed of the rat centered on the time of presentation of the cue. The x-axis is time in seconds with 0 being the time of cue presentation (audio tone) and the y-axis shows the running speed in pixels/second (since a camera was used to track the rats). When a rat started training it didn't discriminate between the cues in a coherent manner. However this changed after about 3 weeks of training for most of the rats (some of the rats took longer, approximately 4 weeks).

Figure 4.1 shows the averaged behavioral response of all rats to the 1, 3 and 5 pellet cues in the value block of the task for a total of 2029 trials or laps. This includes laps completed towards

the end of training and also during the recording sessions. Individual responses of rat R014, R011, R016 and R018 are included as well. Although individual rats showed some variation in their behavioral response to value cues, on average a rat ran fastest for 5 pellets and slowest for 1 pellet. Each individual rat always ran fastest for 5 pellets and slowest for 1 pellet. Thus, the behavior of rats varied depending on the value of the reward predicted by an audio cue.
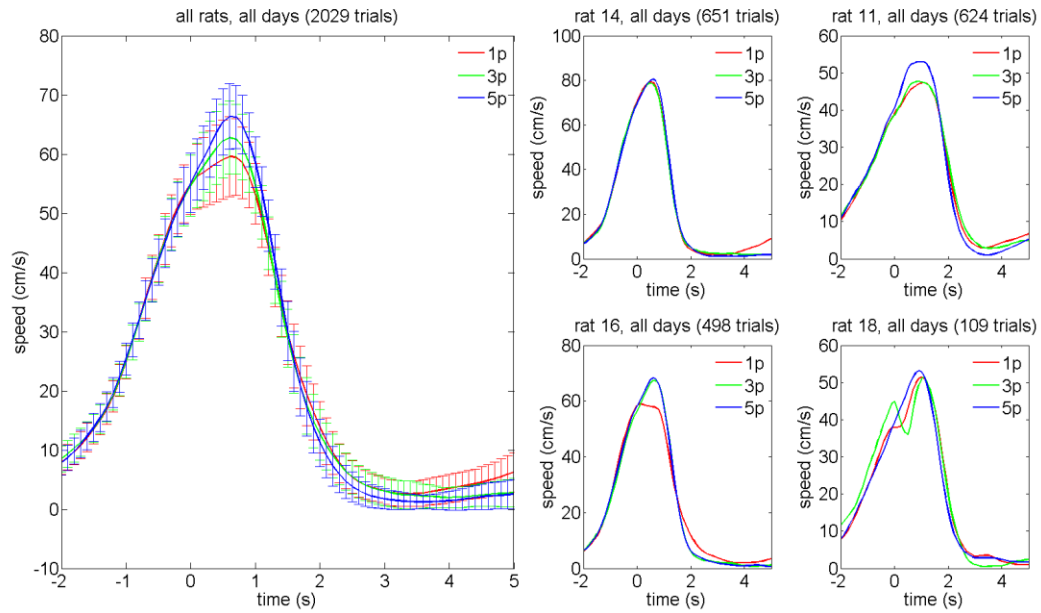

Figure 4.1: Running speed of rats varies in response to different value cues.

Figure 4.1 above shows the running speed of rats in response to different value cues. The left panel shows the average running speed of all rats in response to 1, 3 and 5 pellet cues with error bars. The figures on the right show individual responses of each rat in response to cues. All rats ran faster for a 5 pellet compared to a 1 pellet cue. The variation in running speed is maximum for Rat 16 and minimum for Rat 14. Rat 18 initially slows down and then increases his speed in response to the cue. This is because Rat 18 slowed down on hearing the 3 pellet cue during initial stages of training. The average response for Rat 18 shows a dip in his running speed reflecting the above behavior.

Figure 4.2 is a normalized figure (speed was normalized for each rat by z-scoring it using mean speed and standard deviation) of the above plot to account for differences in average running speed between different rats.
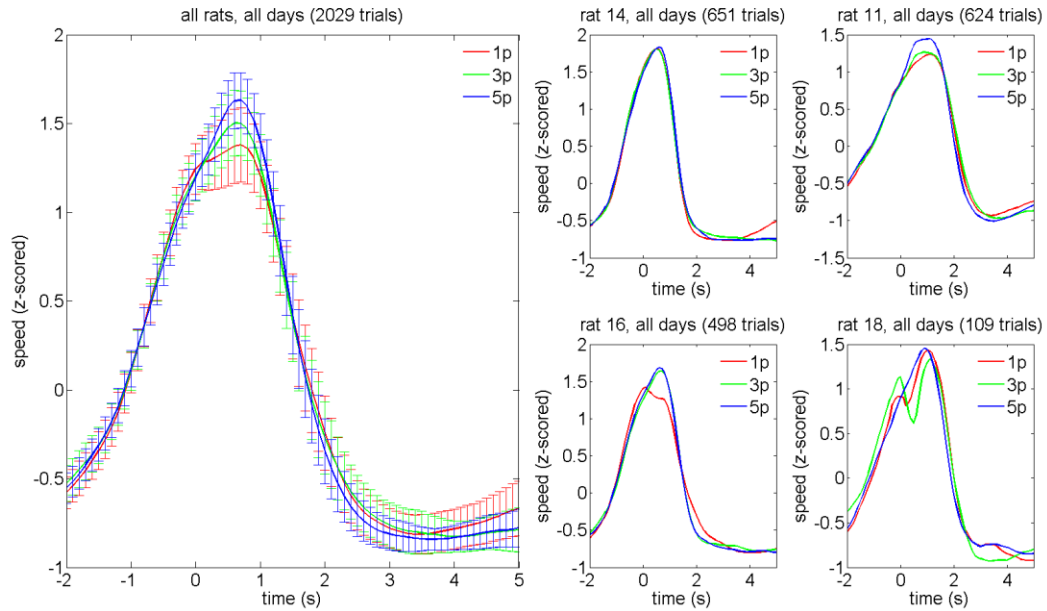
Figure 4.2: Normalized running speed of rats varies in response to different value cues.

The results hold up after normalization. All rats distinguish between 1 and 5 pellet cues. Again the variation in speed is minimum for rat 14 and maximum for rat 16.

4.3 Rats discriminate between cues predicting different risks

The below figures shows the averaged behavioral response of all rats to the low, high and no risk cues in the risk block of the task for a total of 1806 trials or laps. Individual responses of rat R014, R011, R016 and R018 are included in the figure as well. Although individual rats showed some variation in their behavioral response to risk cues, on average rats ran fastest for no risk cue and slowest for the high risk cue.
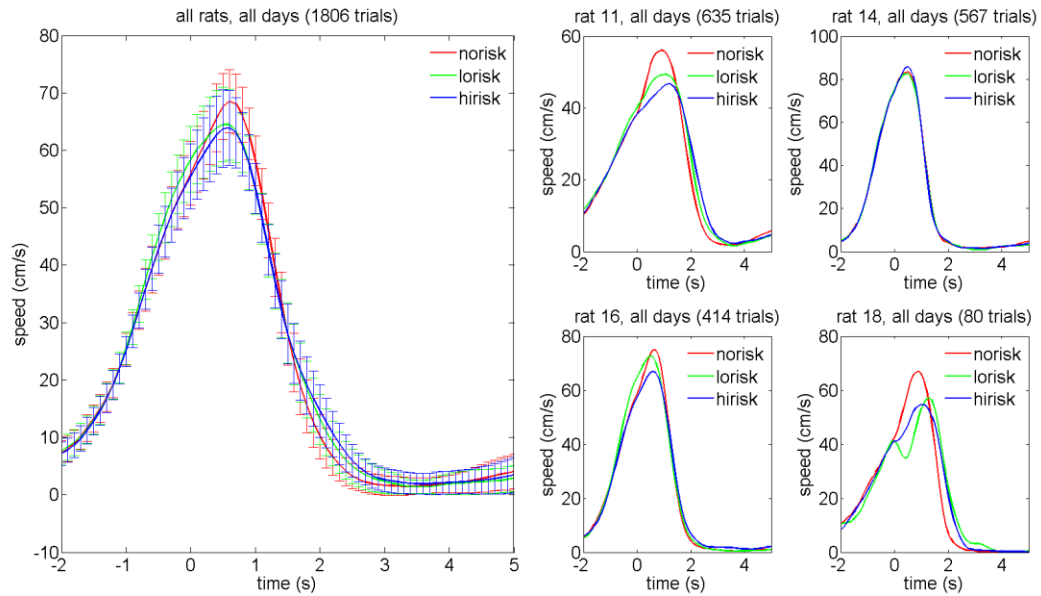
Figure 4.3: Running speed of rats varies in response to different risk cues.

Just like the value task, rats exhibit different running speeds in response to different risk cues. The left panel in figure 4.3 shows the average running speed of all rats in response to no risk, low risk and high risk cues along with error bars. The figures on the right show individual responses of each rat in response to the cues. All rats run faster for no risk cue compared to a low or high cue. The variation in running speed is maximum for Rat 11 and minimum for Rat 14. Rat 18 initially slows down and then increases his speed in response to the no risk cue. This is because Rat 18 initially didn't like the no risk cue (same as the 3 pellet cue from the value task) but got used to it eventually. The average response shows the dip in his running speed reflecting this initial aversion to the no risk cue.

Figure 4.4 below is a normalized version of the above plot to account for variability in running speed.
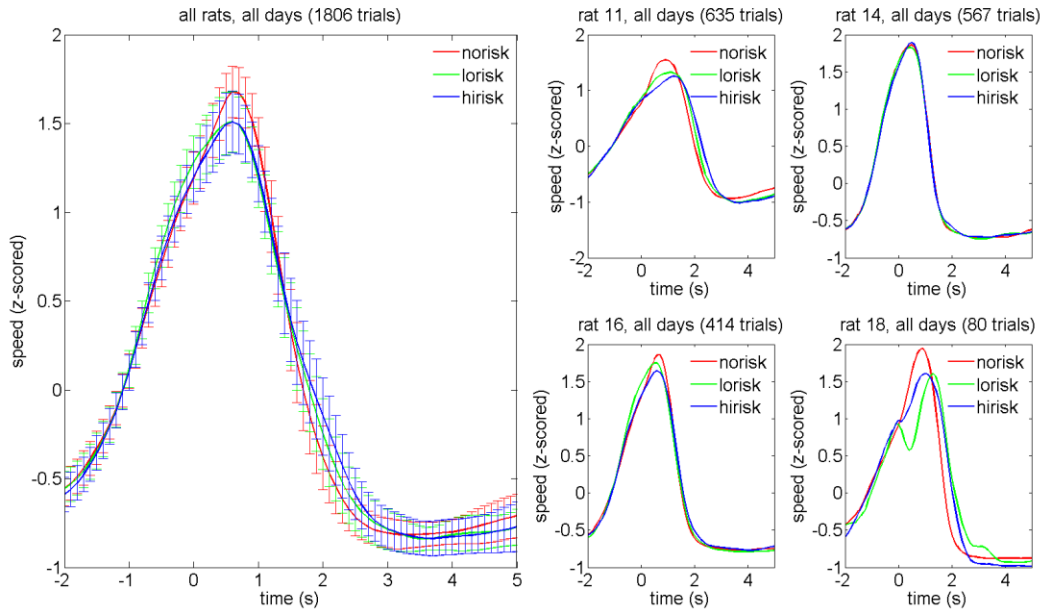
Figure 4.4: Running speed of rats varies in response to different cues.

The average speed plot on the left shows that rats run faster in response to cue associated with no risk compared to cues associated with high and low risk. As before, all the cues are counterbalanced across rats so ensure that the above behavioral preferences are not due to preference or aversion for a particular cue.

4.4 Ventral striatal spiking is modulated by task

Once a rat began to reliably discriminate between different value cues (running faster for the 5 pellet cue than the 1 pellet cue) it was implanted with a vStr electrode array. The electrode array was used to record both extracellular action potentials from neurons and LFP oscillations in the vStr.

The tetrode recording locations for rats R011, R016, R018 and R020 were verified to be in the vStr via histology while R014's tetrodes did not reach the vStr. Any tetrode that missed vStr was excluded from the analysis. Figure 4.5 below shows the recording locations for various rats identified using histology. Only those tetrodes that were verified to be in the vStr are indicated.
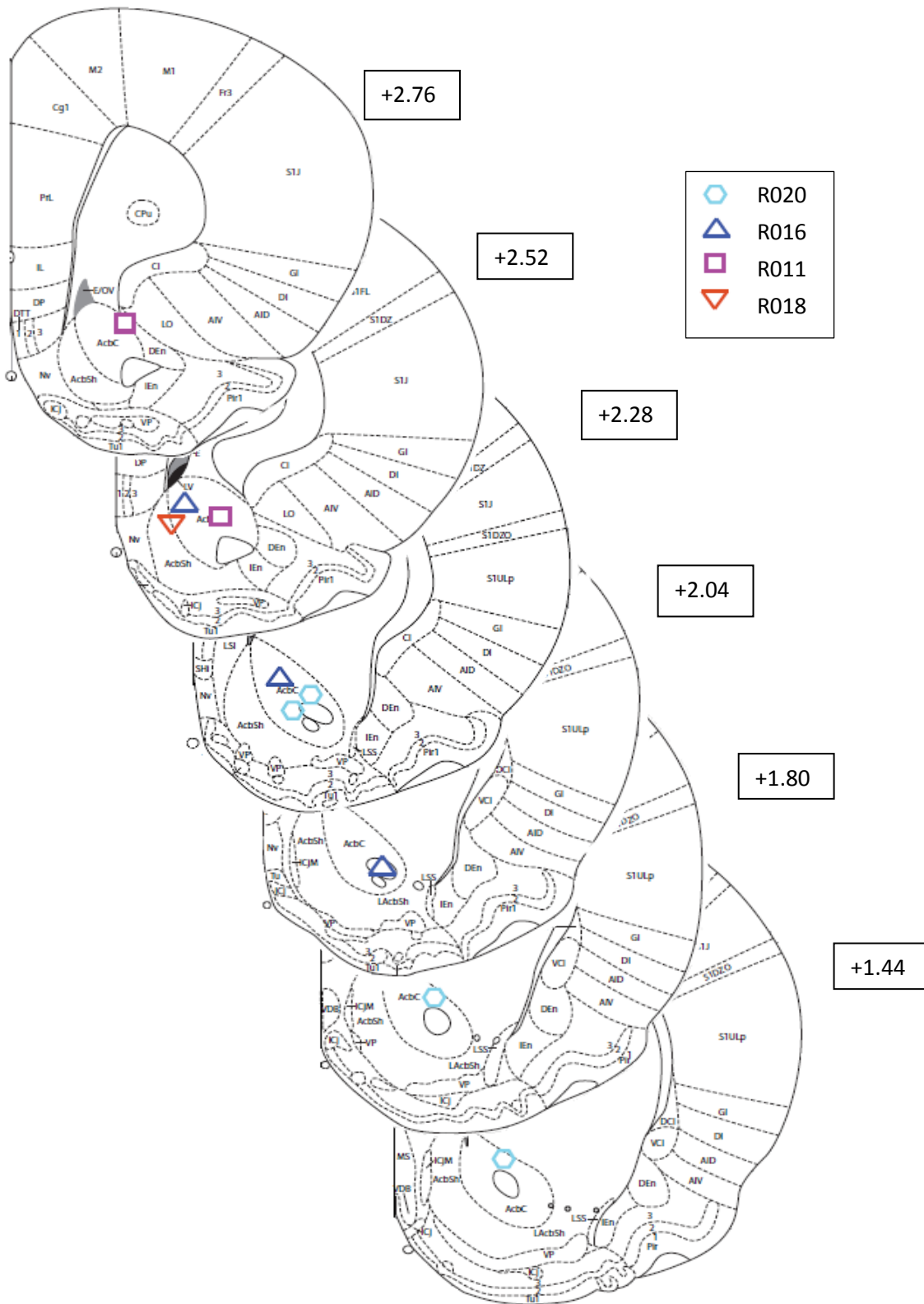
Figure 4.5: Tetrode recording locations. The above figure indicates approximate recording locations identified during histology. The different figures on the brain sections are the recording sites. Different cross-sections are relative to bregma. Rats are

coded by color. All rats were implanted in the right hemisphere. The above figure shows only the tetrodes that were verified to be in the vStr. The brain sections were stained with metachromatic thionin. An example of a stained brain section for R020 is shown below.
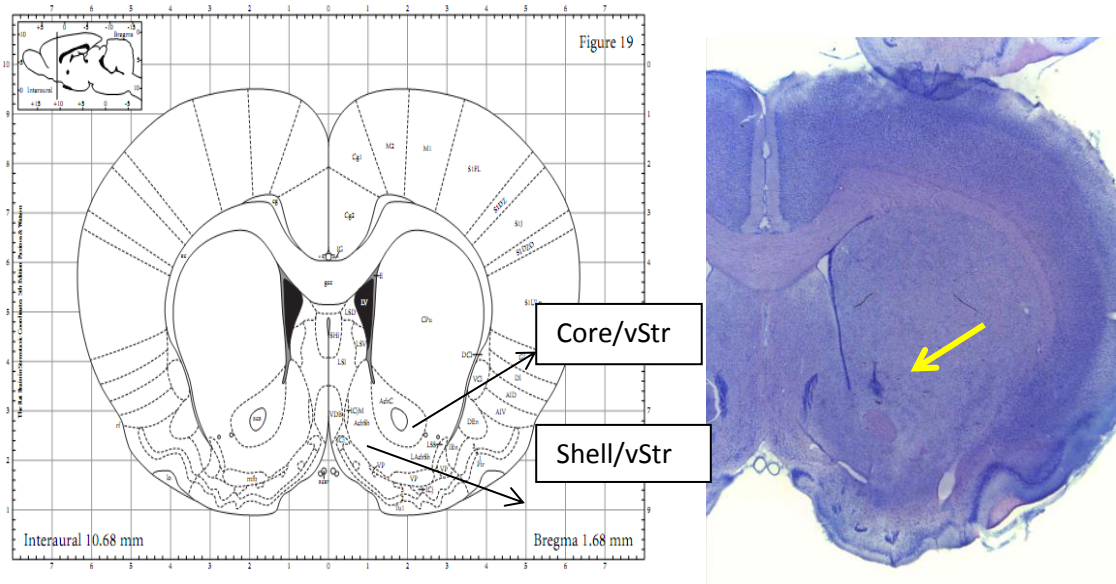
Figure 4.6: Image on the right is a stained coronal section of Rat 20's brain after histology. The tetrode track is shown by the arrow. Image on the left shows the relative position of vStr (nucleus accumbens) in a rat's brain (reprinted from rat atlas).

A total of approximately 20 vStr neurons were recorded during both the value and risk block. About half of the vStr neurons showed a discernible change in their activity with respect to cues and rewards. The number of neurons recorded was too small to do any analysis of the ensemble activity. However a few neurons showed task modulated activity in line with previous recording studies of the vStr (Taha et al., 2006; Taha et al., 2007). Some of the individual neural responses are outlined below:

Below is a plot of vStr neuron recorded from R016 which shows task modulated activity during a particular value block. The first row shows raster plots aligned to the time of cue presentation (red line); the second row includes heat maps showing the change in the firing rate of the neuron over 0.2 second bins; the third row plots the average firing rate of the neuron across all laps during the value block.

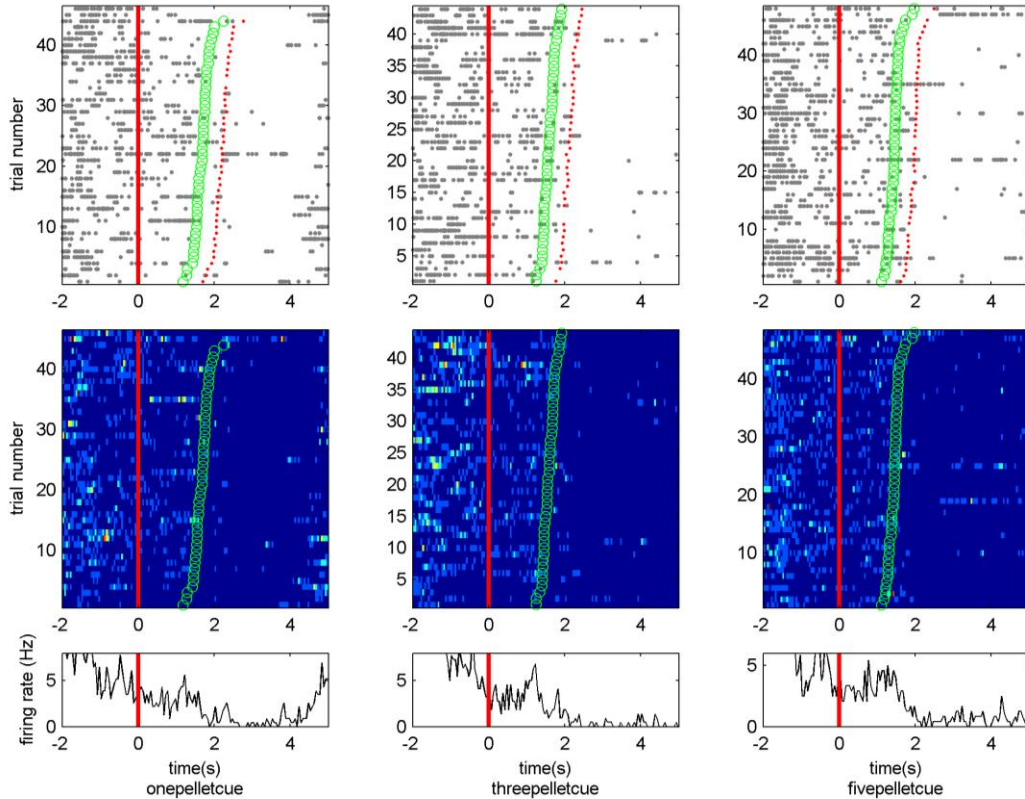Figure 4.7: Response of an example vStr neuron during value block (see text for details)

Figure 4.7 shows the response of an example R016 vStr neuron during the value block. The first column is the neuron's response to the one pellet cue, the second column to the three pellet cue and the last column to the five pellet cue. The top panel shows PETH raster plots of the neuron where the x-axis is time and the y-axis is lap number. The plots are centered on the time of cue presentation shown by the red line. The green circles show the time of nose poke while the red dots show the time of reward arrival. The cyan dots are the time the rat unpokes (first unpoke after nosepoking). The trials are sorted based on the time taken by the rat to nose poke (with the longest time to nose poke on the top and the smallest at the bottom). The time taken by the rat to reach the reward receptacle is smaller for 5 pellets than for 1 pellet Also the above figure combines left and right receptacle trials. The second panel shows the heat maps for the raster plots in the top panel. This is to provide a better visualization of the change in the firing rate of the neuron on a particular lap. The final panel shows the average firing rate of the neuron across all laps. The neuron shows a depression in activity after the arrival of reward. This could possibly be the inhibitory effect of reward consumption on the firing rate of vStr neurons as reported by past studies (Taha et al., 2007).
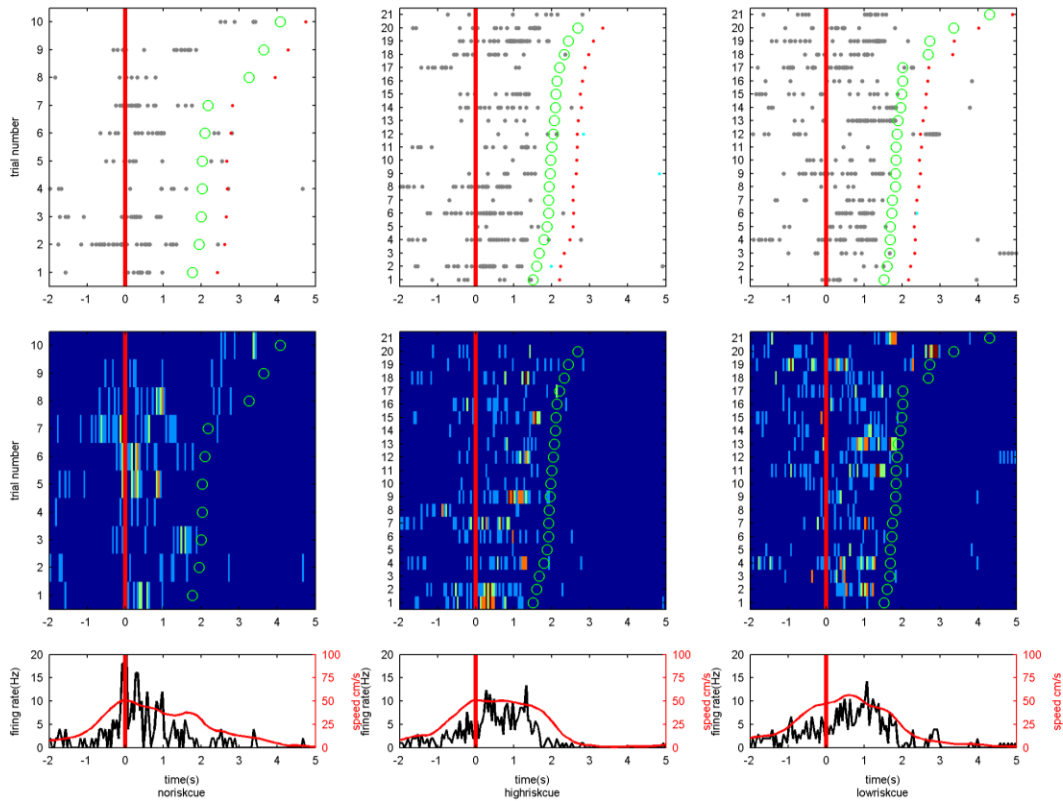
Figure 4.8: Response of an example vStr neuron during risk block (see text for details)

Figure 4.8 shows the activity of a neuron recorded from the vStr of R020 during the risk task. The figure outline is the same as Figure 4.6, however the three columns in this figure represent the response of the neuron to the no risk, high risk and low risk cue respectively. The bold red line shows the time of onset of each of the cues. The thin red line in the last row is the running speed. This neuron seems to increase its firing rate before the arrival of the cue. The firing keeps increasing until reward arrival. Upon receipt of reward the firing rate drops to zero. Thus this neuron seems to show cue and reward modulated activity.

The number of vStr neurons recorded in my task was not enough to analyze ensemble neural activity and its relationship to behavior, cues and rewards.

4.5 Ventral striatum gamma oscillations are modulated by cues and rewards

Along with neurons, local field oscillations were recorded from the vStr of rats R014, R016, R018 and R020. To confirm that low and high gamma were actually present on the tetrodes as

found in previous work (van der Meer and Redish, 2009), Power spectral density (PSD) plots were created for all tetrodes used for LFP analysis. These plots were compared with a reference tetrode in the corpus callosum which a bundle of neural fibers under the cortex having no gamma oscillations (van der Meer & Redish, 2009). A tetrode was used for analysis only if a peak was seen in the gamma band (van der Meer & Redish, 2009). Figure 4.9 below shows the power over various frequencies for two tetrodes in the vStr and a tetrode in the hippocampus of R016.

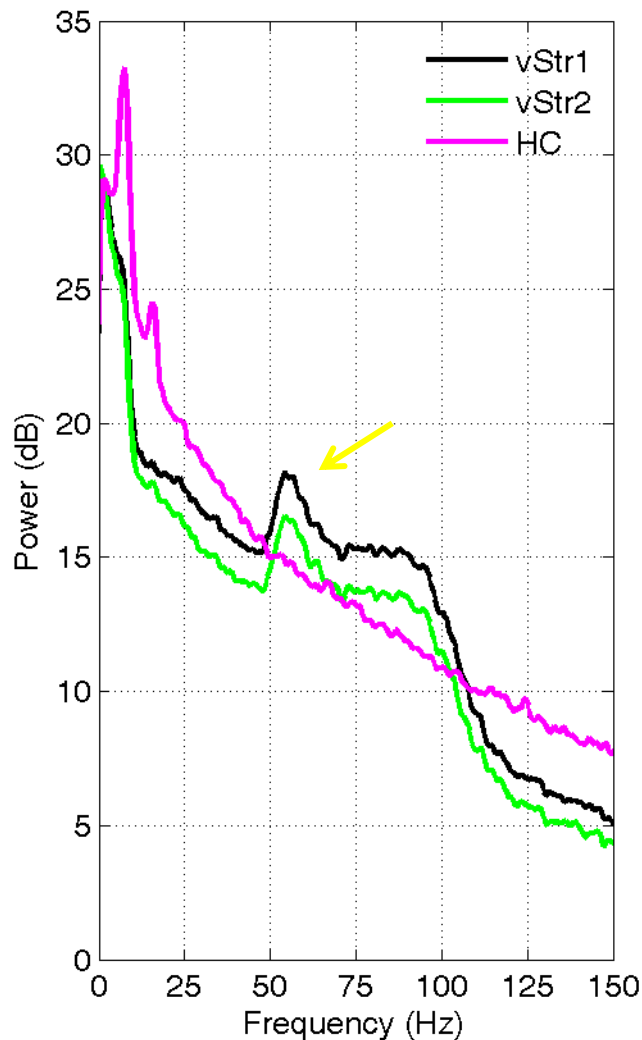Figure 4.9: Power spectral density plot for example tetrodes in the vStr of rat R016

In figure 4.9, the x-axis represents frequencies (in Hz) while the y-axis shows power at a certain frequency. There is clear bump in the power in gamma range for the two vStr tetrodes as indicated by the yellow arrow. However for the hippocampus tetrode, there is no increase in power in the gamma range (40-80 Hz).

Once it was confirmed that a particular tetrode was in vStr (via histology) and had gamma (using the PSD plots), spectrograms were created showing task modulated change in gamma power over both time and space. Time spectrograms were aligned to salient events namely cue presentation, nose poke to obtain reward and the arrival of reward itself.

Below are spectrograms for R020 which show change in gamma power as a function of time in the value task. The spectrograms in the first, second, third rows are centered on cue time, nose poke time and reward time respectively. Also the first, second and third columns represent responses to one pellet, three pellet and five pellet cues respectively. Gamma power seems to increase before the arrival of the cue and also before reward delivery.
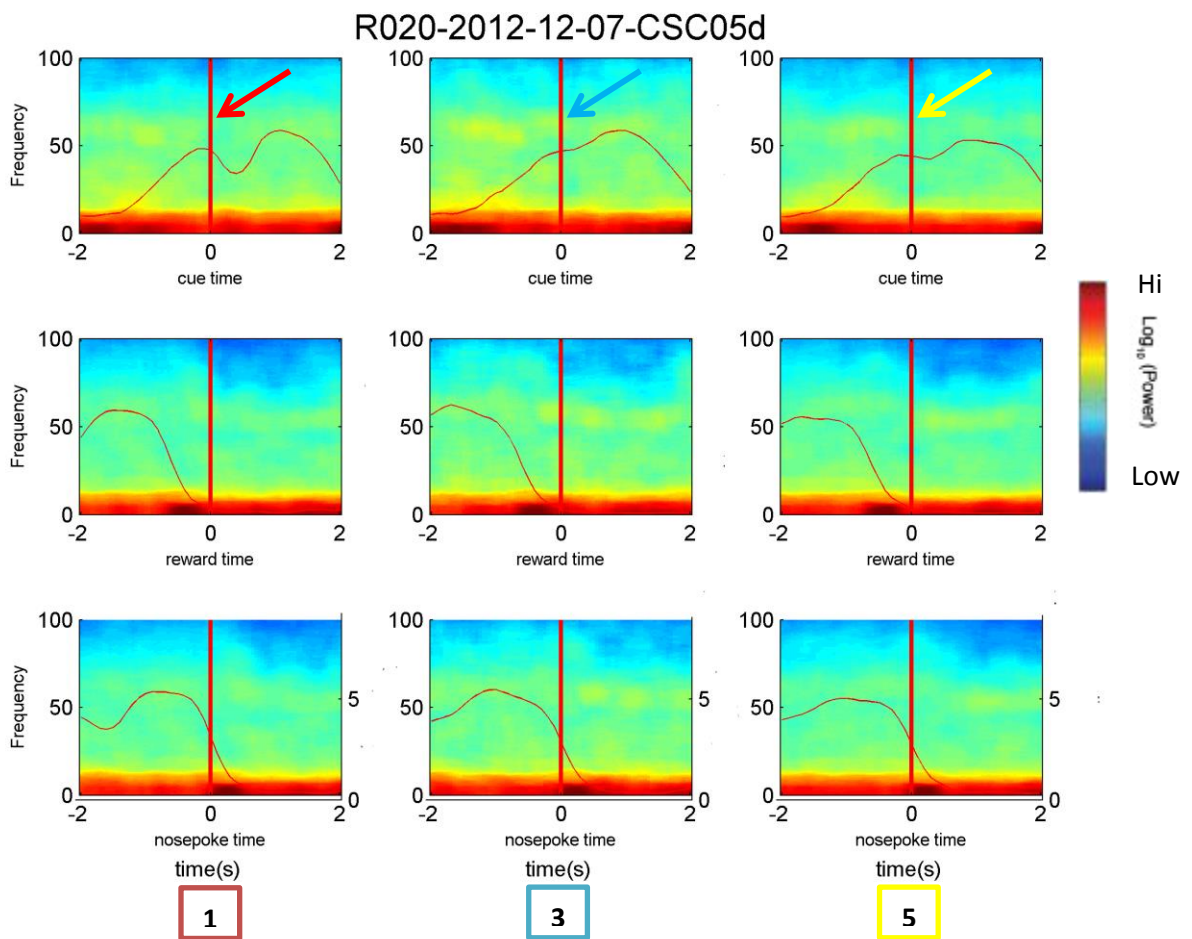
Figure 4.10: Spectrograms showing change in vStr gamma power for R020

Spectrograms in figure 4.10 show change in gamma-50 power over time for Rat 20 while it ran on the value task. The y-axis for all spectrograms is frequency (Hz) while the x-axis is time (seconds). The spectrograms are centered on cue time, nose poke time and reward time (shown by the red line). The first, second and third columns represent responses to one pellet, three

pellet and five pellet cues respectively. The bold red line indicates the time of cue arrival while the thin red line is the speed of the rat centered on the onset of cue. The red, blue and yellow arrows show cue times for one, three and five pellets cues respectively. Also the red, blue and yellow boxes at the bottom show the reward outcome (number of pellets) associated with a particular column. The color of the specific portion of the spectrogram indicates the power of a certain frequency band at a particular time during the task. There is an increase in gamma-50 power before the arrival of the cue and also before the arrival of the reward.
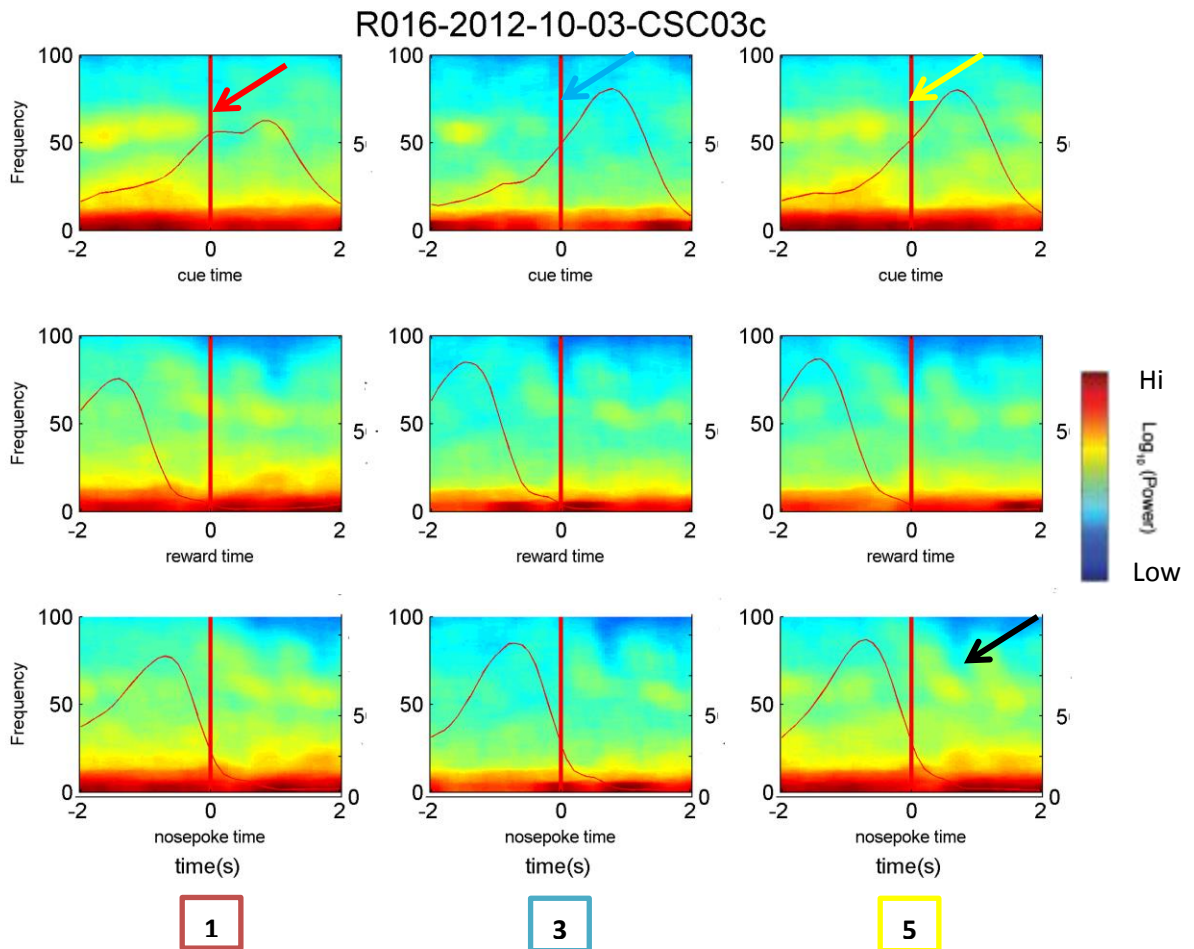
Figure 4.11: Spectrograms showing change in vStr gamma power for R016

Figure 4.11 is the same as the last plot but for an electrode in the vStr of R016. The spectrogram gamma plots for rats R020 and R016 had a transient switch from gamma-80 to gamma-50 around reward delivery as shown in previous studies (van der Meer and Redish, 2009; Howe et al. 2011). An example, gamma-80 to gamma-50 switch is indicated by the black arrow. The switch occurs both before and after reward delivery indicating it might be related to

both anticipation and receipt of reward. However, it is not possible to visibly distinguish between the cues in some systematic way using change in gamma power as a measure. The red, blue and yellow arrows show cue times for one, three and five pellet cues respectively. Also the red, blue and yellow boxes at the bottom show the reward outcome (number of pellets) associated with a particular column.

Spectrograms were also made over space as in previous papers (van der Meer and Redish, 2009) to analyze how gamma power changed as a particular rat approached a reward receptacle. As reported by previous studies, gamma-80 ramped to the reward site while the increase in gamma-50 power was more abrupt and occurred when rat was in close proximity to the reward site (van der Meer and Redish, 2009). The two spectrograms below show change in gamma power as a function of space for R020 during the value task. The x-axis is space (along the length of the track, width of the track is ignored) while the y-axis is power at different frequencies from 0-200 Hz. The first spectrogram shows the change in gamma power as the rat R020 runs from left to right while the second spectrogram shows change in gamma power when rat runs in the opposite direction from right to left.
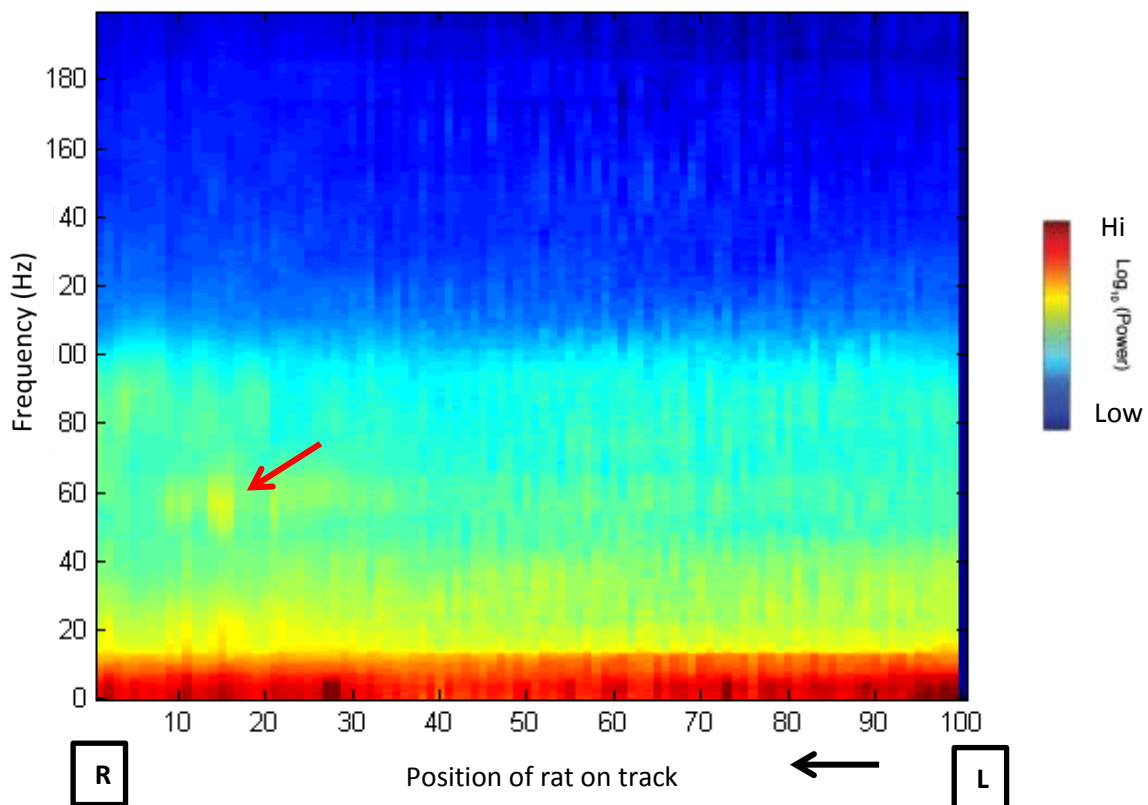
Figure 4.12: Spectrogram showing change in gamma power as function of space

Figure 4.12 shows the change in gamma power (Gamma-50 and Gamma-80) as the rat runs from left feeder (position 100) of the track to the right feeder (position 0) for a five pellet cue on the value task. The y-axis is the frequency while the x-axis the position of the rat on the track. The track is binned into 100 positions namely from 0 to 100. The direction of running is indicated by the black arrow. The color of certain region of the plot indicates the power of a certain frequency at a specific location on the track. Gamma-50 increases abruptly when the rat is in close proximity of the right feeder indicated by a red arrow. The abrupt increase in Gamma-50 around reward sites in line with previous results (van der Meer and Redish, 2009). The times when the rat was idle on the track were excluded while plotting the above spectrogram.
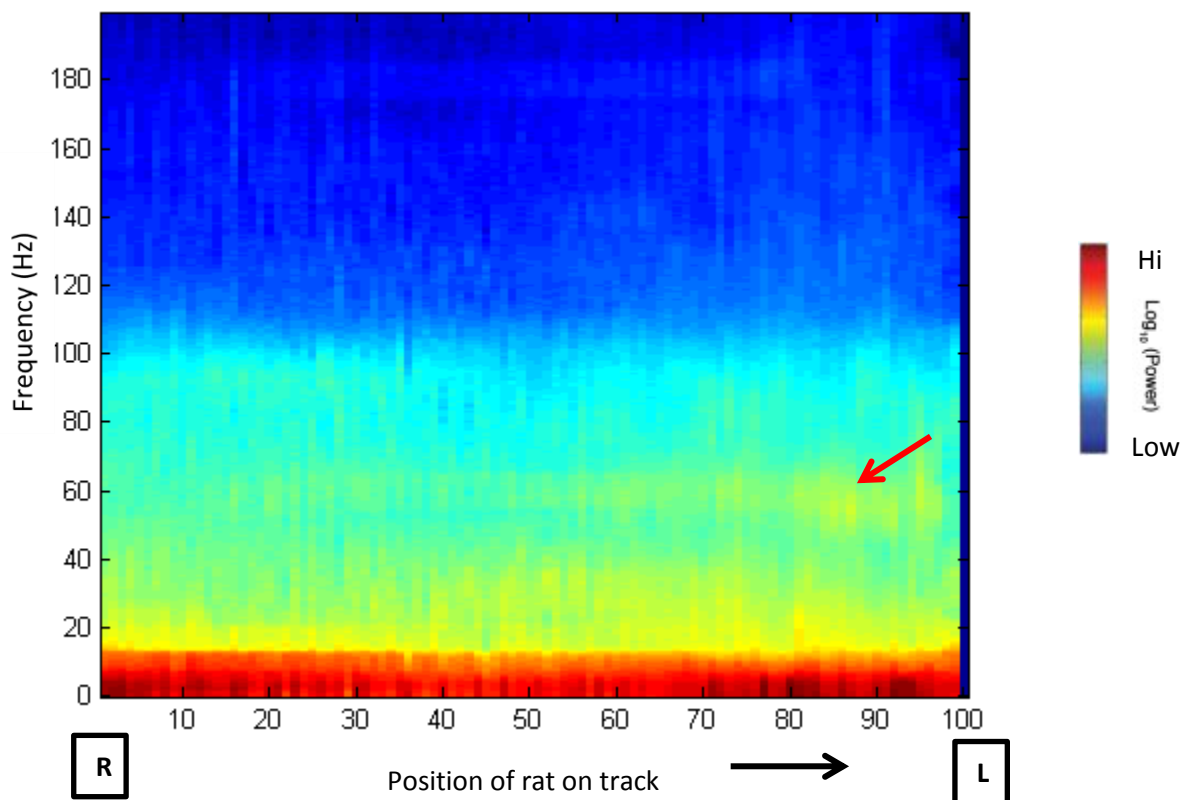
Figure 4.13: Spectrograms showing change in gamma power as function of space

Figure 4.13 shows the change in gamma power (gamma-50 and gamma-80) as the rat runs from right feeder (position 0) on the track to the left feeder (position 100) for a five pellet cue on the value task. The direction of running is indicated by a black arrow. Again, gamma-50 increases as the rat is in close proximity of the left feeder indicated by a red arrow. The times when the rat was idle on the track were excluded while plotting the above spectrogram. Thus gamma-50 and gamma-80 seem to be aligned to rewards in the current experiment. This is discussed further in the next section.

Chapter 5: Conclusion and discussion

Conclusion

The work presented in this thesis shows that the behavior of rats is modulated by value and risk associated with a reward outcome. Specifically, rats can learn about the relationship between cues and the value and risk predicted by the cues with respect to a reward. Ventral striatal gamma oscillations in the current task seem to be modulated by reward delivery in line with previous studies (Howe et al., 2011; Berke et al., 2009; van der Meer and Redish, 2009; Kalenscher et al., 2010). Moreover, there is some evidence that gamma oscillations occur both before the arrival of cue and during the anticipation of reward and receipt of reward.

Behavior

Ventral striatal reward motivated cue tasks generally use at most two or three audio cues to dissociate between different reward outcomes (Roitman et al., 2005). Thus, my task is the first one to my knowledge to successfully train rats on both a value and risk task with five audio cues. My task design allows me to establish if a particular rat learns the association between cues and reward outcomes and also has a controlled nose poke period to disentangle effects of motion artifacts on neural responses. Results on my task show that rats seem to respond in an economically sound fashion on the value task as they run faster for more reward and run faster for non-risky outcomes compared to risky ones. My experiment was done with a group of six rats. Further analysis would possibly enable us to see if there is any interaction between value and risk preferences on a behavior level. It would also help us to identify neural processes that might be involved in processing both value and risk associated with reward outcomes.

Neural recordings

As outlined in the results section, I recorded both vStr neurons and LFP's in my task. However the number of neurons recorded was fairly low which prevented me from doing any ensemble level analysis with the collected data or to determine how single neuron activity is modulated by task variables such as risk. However, I did find vStr neurons that modulated their firing rate with respect to cues and rewards in line with previous studies (Carelli et al., 2000). My analysis focused on identifying the effects of rewards and cues on gamma oscillations. I have included a brief review of previous findings relevant to interpreting findings from my task.

Gamma oscillations in the ventral striatum

Local field oscillations such as alpha, beta, theta and gamma are ubiquitous in a variety of structures in the brain (Schultz et al., 2007, Berens et al., 2010). These oscillations have been investigated by a number of studies and are thought to aid in various sensory, motor and even cognitive processes such as memory and decision making.

Various vStr recording studies have reported gamma oscillations with two distinct bands namely gamma-50 and gamma-80 (Berke et al., 2009; van der Meer and Redish, 2009; Kalenscher et al., 2010). Moreover it has been shown that vStr neurons phase lock to gamma LFP, indicating that these rhythms are related to vStr neural activity and are not just volume conducted from a nearby structure such as the piriform cortex (Berke, 2009; van der Meer and Redish, 2009; Kalenscher et al., 2010). The possible contributors to vStr gamma include excitatory synaptic inputs from afferent structures such as hippocampus, amygdala, prefrontal cortex and thalamus, all of which exhibit gamma oscillations (Colgin et al., 2009; Popescu et al., 2009); local synaptic currents generated by FSI's and MSN's (vStr FSI's show intrinsic bursting in the gamma range (Taverna et al., 2007)) and sub-threshold voltage fluctuations. Volume conduction might be another contributor as Berke et al. (2009) found that vStr gamma-50 is coherent with piriform cortex gamma while gamma-80 is similar to cortical gamma. Thus vStr gamma might be a result of an interaction between external and local neural mechanisms. For example, excitatory inputs from afferent structures might result in excitation of FSI's which would oscillate with gamma frequency and cause widespread inhibition of MSN's through excitatory-inhibitory oscillatory loops. Further experiments (possibly using optogenetics to modulate FSI firing in systematic manner both in vivo/vitro) are required to delineate the exact contributions of the aforementioned sources to vStr gamma.

Both gamma-50 and gamma-80 were clearly visible in my vStr recordings and showed distinct relationships to arrival of cues and rewards. Time spectrograms of vStr gamma show that there is an increase in gamma-50 power before arrival of the cue. Also there is a switch from gamma-80 to gamma-50 after reward delivery as presented in previous studies (van der Meer and Redish, 2009; Howe et al. 2011). However, the above switch also occurs while the rat is anticipating the reward. Thus gamma oscillations might be involved in modulating behavior both before and after receipt of reward. Space spectrograms of gamma power show an abrupt increase in gamma-50 power close to reward delivery while gamma-80 seems to gradually ramp up to both cue and reward in line with previous studies (van der Meer and Redish, 2009).

Modulation of gamma by rewards

Previous recording studies report modulation of both gamma-50 and gamma-80 by reward related events. Specifically, van der Meer and Redish (2009) reported gamma-80 ramping up to the reward site on a T-maze task while gamma-50 increasing abruptly at the reward site. Moreover gamma-80 power fell upon reward receipt and then increased again before returning to baseline. In contrast, Berke and colleagues (2009) found that gamma-50 was completely abolished upon reward receipt while gamma-80 only showed a transient increase to reward delivery. Kalensher et al. also reported an increase in gamma-50 on rewarded trials and a transient increase in gamma-80 before reward delivery. Howe et al. (2011) showed that gamma-50 power was higher on rewarded trials but gamma-80 power was high both before and after goal reaching (the increase in gamma-80 was present initially during learning but diminished significantly as performance on the task improved). Van der Meer and Redish (2009) also analyzed task-modulation of gamma besides rewards. They found that gamma-50 increased transiently before movement initiation but not gamma-80. Also gamma-50 was absent early on during the task and became prominent once the animals learned which arm of the T-maze was rewarded while gamma-80 followed the opposite pattern. Considering the above evidence, the modulation of vStr gamma by rewards and movement is still a topic of active research and is not explained by one normative theory. In my task too, as reported by previous studies, gamma-50 increased abruptly as the rat was in close proximity of a reward. However, the gradual increase of gamma-80 to reward sites was not seen. This might be due to different task structure or result of different techniques used to analyze the data. There was also a shift from gamma-80 to gamma-50 both before and after reward delivery (decrease in gamma-80 power followed by increase in gamma-50 power) which might represent some shift in the neural processing of rewards in the vStr.

Gamma oscillations and decision variables

Although gamma activity shows complex relationships to reward, one of the open questions in the literature is what effect if any do salient aspects of a rewarding outcome such as value and risk have on gamma oscillations. The rats in my project were trained on separate value and risk tasks. They learned to distinguish between cues in an economically sound fashion (based on the value of reward predicted by the cues). Thus, my experiment would be ideal to test if and how gamma oscillations are modulated by reward and risk once learning is complete. It is not clear if vStr gamma would show a systematic change in gamma power in relation to value or risk. If it does, that would be a novel finding informing us about the neural mechanisms that underlie processing of value and risk in the vStr. If it doesn't, that might reflect the fact that my study involves no recording during learning. Previous investigations of vStr mediated behavior

in response to rewards report change in gamma power over the course of learning. Gamma power (especially gamma-80) is high initially but seems to become constant/diminishes once a certain learning criterion has been met (Howe 2011, van der Meer and Redish, 2009). Also the vStr as mentioned in earlier chapters gets substantial projections from dopamine neurons in the VTA. Current theories of vStr dopamine postulate its role in invigorating behavior in response to rewards and reward predictive cues (Berridge and Robinson, 1998; Salamone et al., 2007; Nicola, 2007). Injection of dopamine receptor antagonists into the vStr specifically reduce an animal's ability to exhibit flexible behaviors in order to obtain rewards (Nicola, 2007). Moreover, both reward predictive cues and rewards cause phasic elevations in vStr dopamine particularly in tasks involving flexible approach behavior (Roitman et al., 2004, Day et al., 2010). Rats on my task had to exhibit the same behavior or action to collect rewards on each trial. Thus, it is possible that vStr dopamine might have minimally contributed to the decisions made (or actions taken) by a rat in my task especially once learning was complete and the actions executed became more of a habit possibly mediated by the dorsal striatum (Atallah et al., 2007). Since vStr dopamine not only changes the activity dynamics of single vStr neurons but also of gamma oscillations (Berke et al., 2009), it might be possible that no real change would be seen in the power of gamma oscillations on my task due to a lack of flexible actions. It would be interesting to record simultaneously from the dorsal and ventral striatum in my task to see what changes in gamma power take place in both structures as learning progresses.

Future directions

Further studies are required to investigate the precise behavioral correlates of gamma oscillations especially during learning. Ventral striatum receives a confluence of inputs from a number of afferent structures. It also receives dopaminergic input from the VTA which modulates the activity of vStr neurons in a complicated manner through various neural pathways (Seasack and Grace, 2010). Simultaneous recording of ensembles of vStr neurons along with novel techniques such as opto-genetics and voltammetry will enable us to examine the role of vStr neural activity in the processing of cues and rewards. A multifaceted approach is necessary to disentangle the various neural mechanisms implicated in reward-motivated behavior. Finally, controlled experiments examining the effects of decision variables such as value and risk on behavior and the underlying neural activity might enable us to better understand how we make complicated decisions, which is a big part of what makes us human.

Bibliography

Apicella, P., Scarnati, E., and Schultz, W. (1991). Tonically discharging neurons of monkey striatum respond to preparatory and rewarding stimuli. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, *84*(3), 672–5.

Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., and O'Reilly, R. C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nature neuroscience*, *10*(1), 126–31.

Berke, J D. (2009). Fast oscillations in cortical-striatal networks switch frequency following rewarding events and stimulant drugs. *The European journal of neuroscience*, *30*(5), 848–59.

Berke, Joshua D, Okatan, M., Skurski, J., and Eichenbaum, H. B. (2004). Oscillatory entrainment of striatal neurons in freely moving rats. *Neuron*, *43*(6), 883–96.

Berridge, K. C., and Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain research. Brain research reviews*, *28*(3), 309–69.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*, *191*(3), 391–431.

Bracci, E., Centonze, D., Bernardi, G., and Calabresi, P. (2003). Voltage-dependent membrane potential oscillations of rat striatal fast-spiking interneurons. *The Journal of physiology*, *549*(Pt 1), 121–30.

Brown, E. N., Kass, R. E., and Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature neuroscience*, *7*(5), 456–61.

Buzsáki, G. (2002). Theta oscillations in the hippocampus. *Neuron*, *33*(3), 325–40.

Cardinal, R N, Pennicott, D. R., Sugathapala, C. L., Robbins, T. W., and Everitt, B. J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science (New York, N.Y.)*, *292*(5526), 2499–501.

Cardinal, Rudolf N, Parkinson, J. a, Hall, J., and Everitt, B. J. (2002a). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and biobehavioral reviews*, *26*(3), 321–52.

Carelli, R. M., and Deadwyler, S. a. (1994). A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *14*(12), 7735–46.

Carelli, R. M., Ijames, S. G., and Crumling, A. J. (2000). Evidence that separate neural circuits in the nucleus accumbens encode cocaine versus "natural" (water and food) reward. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *20*(11), 4255–66.

Carelli, R. M. (2004). Nucleus accumbens cell firing and rapid dopamine signaling during goal-directed behaviors in rats. *Neuropharmacology*, *47*, 180–189.

Chapter 2 : Introduction to Point Processes. (2003), 1–40.

Cohen, M. X., Axmacher, N., Lenartz, D., Elger, C. E., Sturm, V., and Schlaepfer, T. E. (2009a). Good vibrations: cross-frequency coupling in the human nucleus accumbens during reward processing. *Journal of cognitive neuroscience*, *21*(5), 875–89.

Cohen, M. X., Axmacher, N., Lenartz, D., Elger, C. E., Sturm, V., and Schlaepfer, T. E. (2009b). Nuclei accumbens phase synchrony predicts decision-making reversals following negative feedback. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(23), 7591–8.

Corbit, L H, Muir, J. L., & Balleine, B. W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *21*(9), 3251–60.

Corbit, Laura H, & Balleine, B. W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *25*(4), 962–70.

Crespi, Leo P. (1942). Quantitative variation of incentive and performance in the white rat. *American Journal of Psychology, 55,* 467-517

Dalley, J. W., Everitt, B. J., and Robbins, T. W. (2011). Impulsivity, compulsivity, and top-down cognitive control. *Neuron*, *69*(4), 680–94.

Day, J. J., Jones, J. L., and Carelli, R. M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *The European journal of neuroscience*, *33*(2), 308–21.

Day, J. J., Jones, J. L., Wightman, R. M., and Carelli, R. M. (2010). Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. *Biological psychiatry*, *68*(3), 306–9.

Delamater, A. R. (1995). Outcome-selective effects of intertrial reinforcement in a Pavlovian appetitive conditioning paradigm with rats. *Animal Learning and Behavior*, *23*(1), 31–39.

Donnell, O., and Grace, A. (1995). Synaptic Interactions Accumbens Neurons : Input among Excitatory Afferents to Nucleus Hippocampal Gating of Prefrontal Cortical, *15*(May), 3622–3639.

Epstein, J., Pan, H., Kocsis, J. H., Yang, Y., Butler, T., Chusid, J., Hochberg, H., et al. (2006). Lack of ventral striatal response to positive stimuli in depressed versus normal subjects. *The American journal of psychiatry*, *163*(10), 1784–90.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, *9*(10), 474–80.

Goldstein, B. L., Barnett, B. R., Vasquez, G., Tobia, S. C., Kashtelyan, V., Burton, A. C., Bryden, D. W., et al. (2012). Ventral striatum encodes past and predicted value independent of motor contingencies. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *32*(6), 2027–36.

Gruber, A. J., Hussain, R. J., and O'Donnell, P. (2009). The nucleus accumbens: a switchboard for goal-directed behaviors. *PloS one*, *4*(4), e5062.

Hariri, A. R., Brown, S. M., Williamson, D. E., Flory, J. D., de Wit, H., and Manuck, S. B. (2006). Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *26*(51), 13213–7.

Harris, K. D., Henze, D. A., Csicsvari, J., Hirase, H., and Buzsaki, G. (2000). Accuracy of Tetrode Spike Separation as Determined by Simultaneous Intracellular and Extracellular Measurements. *J Neurophysiol*, *84*(1), 401–414.

Henze, D. A., Borhegyi, Z., Csicsvari, J., Mamiya, A., Harris, K. D., and Buzsaki, G. (2000). Intracellular Features Predicted by Extracellular Recordings in the Hippocampus In Vivo. *J Neurophysiol*, *84*(1), 390–400.

Hollerman, J. R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *Journal of neurophysiology*, *80*(2), 947–63.

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of experimental psychology. Animal behavior processes*, *30*(2), 104–17.

Howe, M. W., Atallah, H. E., McCool, A., Gibson, D. J., and Graybiel, A. M. (2011). Habit learning is associated with major shifts in frequencies of oscillatory activity and synchronized spike firing in striatum. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(40), 16801–6.

Humphries, M. D., and Prescott, T. J. (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. *Progress in neurobiology*, *90*(4), 385–417.

Johnston, D., & Wu, S. M.-S. (1994). *Foundations of Cellular Neurophysiology (Bradford Books)* (p. 710). A Bradford Book.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, *63*(6), 733–45.

Kandel, E., Schwartz, J., & Jessell, T. (2000). *Principles of Neural Science* (p. 1414). McGraw-Hill Medical.

Katzner, S., Nauhaus, I., Benucci, A., Bonin, V., Ringach, D. L., and Carandini, M. (2009). Local origin of field potentials in visual cortex. *Neuron*, *61*(1), 35–41.

Kalenscher, T., Lansink, C. S., Lankelma, J. V., and Pennartz, C. M. A. (2010). Reward-associated gamma oscillations in ventral striatum are regionally differentiated and modulate local firing activity. *Journal of neurophysiology*, *103*(3), 1658–72.

Khamassi, Mehdi, Mulder, A. B., Tabuchi, E., Douchamps, V., and Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal neurons of behaving rats. *The European journal of neuroscience*, *28*(9), 1849–66.

Knutson, B., Adams, C. M., Fong, G. W., and Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *21*(16), RC159.

Kuhnen, C. M., and Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, *47*(5), 763–70.

Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Joosten, R. N. J. M. a, McNaughton, B. L., and Pennartz, C. M. a. (2008). Preferential reactivation of motivationally relevant information in the ventral striatum. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *28*(25), 6372–82.

Lansink, C. S., Goltstein, P. M., Lankelma, J. V., Mcnaughton, B. L., and Ma, C. (n.d.). Text S1 Hippocampus Leads Ventral Striatum in Replay of Place - Reward Information, 1–15.

Levine, Daniel S., Brown, Vincent R., Timothy Shirey (1999). Oscillations in Neural Systems. *Psychology Press*.

Leventhal, D. K., Gage, G. J., Schmidt, R., Pettibone, J. R., Case, A. C., & Berke, J. D. (2012). Basal ganglia beta oscillations accompany cue utilization. *Neuron*, *73*(3), 523–36.

McGinty, V. B., Lardeux, S., Taha, S. A., Kim, J. J., and Nicola, S. M. (2013). Invigoration of reward seeking by cue and proximity encoding in the nucleus accumbens. *Neuron*, *78*(5), 910–22.

Mitra, P., and Bokil, H. (2007). *Observed Brain Dynamics* (p. 408). Oxford University Press, USA.

Mitra, P. P., and Pesaran, B. (1999). Analysis of dynamic brain imaging data. *Biophysical journal*, *76*(2), 691–708.

Mitzdorf, U. (1985). Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. *Physiol Rev*, *65*(1), 37–100.

Miyazaki, K., Mogi, E., and Araki, N. (1998). Dependent anticipation in rat nucleus accumbens, *9*(17), 3943–3948.

Mizumori, S. J. Y., Yeshenko, O., Gill, K. M., and Davis, D. M. (2004). Parallel processing across neural systems: implications for a multiple memory system hypothesis. *Neurobiology of learning and memory*, *82*(3), 278–98.

Mogenson, G. J., Jones, D. L., and Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Progress in neurobiology*, *14*(2-3), 69–97.

Morra, J. T., Glick, S. D., and Cheer, J. F. (2010). Neural encoding of psychomotor activation in the nucleus accumbens core, but not the shell, requires cannabinoid receptor signaling. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(14), 5102–7.

Mulder, A. B., Shibata, R., Trullier, O., and Wiener, S. I. (2005). Spatially selective reward site responses in tonically active neurons of the nucleus accumbens in behaving rats. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, *163*(1), 32–43.

Münte, T. F., Heldmann, M., Hinrichs, H., Marco-Pallares, J., Krämer, U. M., Sturm, V., and Heinze, H.-J. (2007). Nucleus Accumbens is Involved in Human Action Monitoring: Evidence from Invasive Electrophysiological Recordings. *Frontiers in human neuroscience*, *1*, 11.

Nicola, S. M., Taha, S. A., Kim, S. W., & Fields, H. L. (2005). Nucleus accumbens dopamine release is necessary and sufficient to promote the behavioral response to reward-predictive cues. *Neuroscience*, *135*(4), 1025–33.

Nicola, S. M. (2007). The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology*, *191*(3), 521–50.

Nicola, S. M. (2010). 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. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(49), 16585–600.

O'Neill, M., and Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron*, *68*(4), 789–800.

Parkinson, J A, Willoughby, P. J., Robbins, T. W., and Everitt, B. J. (2000). Disconnection of the anterior cingulate cortex and nucleus accumbens core impairs Pavlovian approach behavior: further evidence for limbic cortical-ventral striatopallidal systems. *Behavioral neuroscience*, *114*(1), 42–63.

Parkinson, J. A., Robbins, T. W., and Everitt, B. J. (n.d.). Selective excitotoxic lesions of the nucleus accumbens core and shell differentially affect aversive Pavlovian conditioning to discrete and contextual cues.

Pennartz, C. M., Groenewegen, H. J., and Lopes da Silva, F. H. (1994). The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomical data. *Progress in neurobiology*, *42*(6), 719–61.

Pennartz, C. M. A., Berke, J. D., Graybiel, A. M., Ito, R., Lansink, C. S., van der Meer, M., Redish, A. D., et al. (2009). Corticostriatal Interactions during Learning, Memory Processing, and Decision Making. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(41), 12831–8.

Pennartz, C. M. A., Ito, R., Verschure, P. F. M. J., Battaglia, F. P., and Robbins, T. W. (2011). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends in neurosciences*, *34*(10), 548–59.

Peters, J., Bromberg, U., Schneider, S., Brassen, S., Menz, M., Banaschewski, T., Conrod, P. J., et al. (2011). Lower ventral striatal activation during reward anticipation in adolescent smokers. *The American journal of psychiatry*, *168*(5), 540–9.

Platt, M. L., and Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nature neuroscience*, *11*(4), 398–403.

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature reviews. Neuroscience*, *9*(7), 545–56.

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black & W.F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99).

Reynolds, S. M., and Berridge, K. C. (2001). Fear and feeding in the nucleus accumbens shell: rostrocaudal segregation of GABA-elicited defensive behavior versus eating behavior. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *21*(9), 3261–70.

Roitman, M. F., Wheeler, R. A., and Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron*, *45*(4), 587–97.

Saddoris, M. P., Stamatakis, A., and Carelli, R. M. (2011). Neural correlates of Pavlovian-to-instrumental transfer in the nucleus accumbens shell are selectively potentiated following cocaine self-administration. *The European journal of neuroscience*, *33*(12), 2274–87.

Salamone, J. D., Correa, M., Farrar, A., and Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, *191*(3), 461–82.

Sanes, J. N., and Donoghue, J. P. (1993). Oscillations in local field potentials of the primate motor cortex during voluntary movement. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(10), 4470–4.

Scheres, A., Milham, M. P., Knutson, B., and Castellanos, F. X. (2007). Ventral striatal hyporesponsiveness during reward anticipation in attention-deficit/hyperactivity disorder. *Biological psychiatry*, *61*(5), 720–4.

Schoenbaum, G., and Setlow, B. (2003). Lesions of nucleus accumbens disrupt learning about aversive outcomes. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *23*(30), 9833–41.

Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K., and Redish, A. D. (2005). Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience*, *131*(1), 1–11.

Schultz, W. (2000). Multiple reward signals in the brain. *Nature reviews. Neuroscience*, *1*(3), 199–207.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *13*(3), 900–13.

Schultz, W., Apicella, P., Scarnati, E., and Ljungberg, T. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *12*(12), 4595–610.

Schultz, W. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599.

Seamans, J. K., and Phillips, A. G. (1994). Selective memory impairments produced by transient lidocaine-induced lesions of the nucleus accumbens in rats. *Behavioral neuroscience*, *108*(3), 456–68.

Sesack, S. R., and Grace, A. A. (2010). Cortico-Basal Ganglia reward network: microcircuitry. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, *35*(1), 27–47.

Shibata, R., Mulder, A. B., Trullier, O., and Wiener, S. I. (2001). Position sensitivity in phasically discharging nucleus accumbens neurons of rats alternating between tasks requiring complementary types of spatial cues. *Neuroscience*, *108*(3), 391–411.

Simmons, J. M., Ravel, S., Shidara, M., and Richmond, B. J. (2007). A comparison of reward-contingent neuronal activity in monkey orbitofrontal cortex and ventral striatum: guiding actions toward rewards. *Annals of the New York Academy of Sciences*, *1121*, 376–94.

Sirota, A., Montgomery, S., Fujisawa, S., Isomura, Y., Zugaro, M., and Buzsáki, G. (2008). Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. *Neuron*, *60*(4), 683–97.

Skinner B F. The behavior of organisms: an experimental analysis. New York: Appleton-Century-Crofts. 1938. 457 p.

Sugam, J. A., Saddoris, M. P., & Carelli, R. M. (2013). Nucleus Accumbens Neurons Track Behavioral Preferences and Reward Outcomes During Risky Decision Making. *Biological Psychiatry*.

Talmi, D., Seymour, B., Dayan, P., and Dolan, R. J. (2008). Human Pavlovian-instrumental transfer. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *28*(2), 360–8.

Taha, S. A., and Fields, H. L. (2005). Encoding of palatability and appetitive behaviors by distinct neuronal populations in the nucleus accumbens. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *25*(5), 1193–202.

Taha, S. A., and Fields, H. L. (2006). Inhibitions of nucleus accumbens neurons encode a gating signal for reward-directed behavior. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *26*(1), 217–22.

Taha, S. A., Nicola, S. M., and Fields, H. L. (2007). Cue-evoked encoding of movement planning and execution in the rat nucleus accumbens. *The Journal of physiology*, *584*(Pt 3), 801–18.

Taverna, S., Canciani, B., and Pennartz, C. M. a. (2007). Membrane properties and synaptic connectivity of fast-spiking interneurons in rat ventral striatum. *Brain research*, *1152*, 49–56.

Tobler, P. N., O'doherty, J. P., Dolan, R. J., and Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of neurophysiology*, *95*(1), 301–10.

Tort, A. B. L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology*, *104*(2), 1195–210.

Tremblay, L., Hollerman, J. R., and Schultz, W. (1998). Modifications of reward expectation-related neuronal activity during learning in primate striatum. *Journal of neurophysiology*, *80*(2), 964–77.

Wyvell, C. L., and Berridge, K. C. (2000). Intra-Accumbens Amphetamine Increases the Conditioned Incentive Salience of Sucrose Reward: Enhancement of Reward "Wanting" without Enhanced "Liking" or Response Reinforcement. *J. Neurosci.*, *20*(21), 8122–8130.

Xia, Y., Driscoll, J. R., Wilbrecht, L., Margolis, E. B., Fields, H. L., and Hjelmstad, G. O. (2011). Nucleus accumbens medium spiny neurons target non-dopaminergic neurons in the ventral tegmental area. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *31*(21), 7811–6.

Zink, C. F., Pagnoni, G., Martin, M. E., Dhamala, M., and Berns, G. S. (2003). Human striatal response to salient nonrewarding stimuli. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *23*(22), 8092–7.

van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C., and Redish, A. D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron*, *67*(1), 25–32.

van der Meer, M. A. A., Kalenscher, T., Lansink, C. S., Pennartz, C. M. A., Berke, J. D., and Redish, A. D. (2010). Integrating early results on ventral striatal gamma oscillations in the rat. *Frontiers in neuroscience*, *4*, 300.

van der Meer, M. A. A., and Redish, A. D. (2009a). Covert Expectation-of-Reward in Rat Ventral Striatum at Decision Points. *Frontiers in integrative neuroscience*, *3*, 1.

van der Meer, M. A. A., and Redish, A. D. (2009). Low and High Gamma Oscillations in Rat Ventral Striatum have Distinct Relationships to Behavior, Reward, and Spiking Activity on a Learned Spatial Decision Task. *Frontiers in integrative neuroscience*, *3*, 9.

van der Meer, M. A. A., and Redish, A. D. (2010). Expectancies in decision making, reinforcement learning, and ventral striatum. *Frontiers in neuroscience*, *4*, 6.

van der Meer, M. A. A., and Redish, A. D. (2011). Theta phase precession in rat ventral striatum links place and reward information. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *31*(8), 2843–54.

Williams, L. E., and Bargh, J. A. (2008, 24 October). Experiencing physical warmth promotes interpersonal warmth. *Science*, 322, 606-607.