Dynamics of dopamine signaling and network activity in the striatum during learning and motivated pursuit of goals

by

Mark W. Howe
B.S. Biology
University of Wisconsin-Madison, 2006

SUBMITTED TO THE DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN NEUROSCIENCE
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2013

Signature of author: _____

Department of Brain and Cognitive Sciences
September 28, 2012

Certified by: _____

Ann M. Graybiel
Walter A. Rosenblith Professor of Neuroscience
Thesis Supervisor

Accepted by: _____

Matthew A. Wilson
Sherman Fairchild Professor of Neuroscience
Director of Graduate Education for Brain and Cognitive Science

Dynamics of dopamine signaling and network activity in the striatum during learning and motivated pursuit of goals

by

Mark W. Howe
B.S. Biology
University of Wisconsin-Madison, 2006

Submitted to the Department of Brain and Cognitive Sciences
on September 28, 2012 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Neuroscience

Learning to direct behaviors towards goals is a central function of all vertebrate nervous systems. Initial learning often involves an exploratory phase, in which actions are flexible and highly variable. With repeated successful experience, behaviors may be guided by cues in the environment that reliably predict the desired outcome, and eventually behaviors can be executed as crystallized action sequences, or "habits", which are relatively inflexible. Parallel circuits through the basal ganglia and their inputs from midbrain dopamine neurons are believed to make critical contributions to these phases of learning and behavioral execution. To explore the neural mechanisms underlying goal-directed learning and behavior, I have employed electrophysiological and electrochemical techniques to measure neural activity and dopamine release in networks of the striatum, the principle input nucleus of the basal ganglia as rats learned to pursue rewards in mazes. The electrophysiological recordings revealed training dependent dynamics in striatum local field potentials and coordinated neural firing that may differentially support both network rigidity and flexibility during pursuit of goals. Electrochemical measurements of real-time dopamine signaling during maze running revealed prolonged signaling changes that may contribute to motivating or guiding behavior. Pathological over or under-expression of these network states may contribute to symptoms experienced in a range of basal ganglia disorders, from Parkinson's disease to drug addiction.

Thesis supervisor: Ann Graybiel
Title: Institute Professor

Table of Contents

Acknowledgements

The work presented in this thesis is the product of a concerted effort of many members of the Graybiel lab and the Phillips lab, who have contributed valuable advice and technical support throughout the process. I am particularly grateful to my advisor Dr. Ann Graybiel, for her unwavering support and encouragement of this project and her scientific feedback and advice on everything from broad conceptual direction to technical approach and data interpretation. I would also like to thank the following people: My officemates Hisham Atallah and Eric Burguiere for the frequent valuable discussions regarding interpretation of my data, and for their candid and insightful critiques of my conclusions. I'm also grateful to Hisham for providing me access to his data (see Chapter 2). Patrick Tierney for helping me to develop and set up the FSCV experiments. Paul Phillips, Stefan Sandberg, and the rest of the Phillips lab for sharing their expertise in FSCV and allowing me to learn the technique in their lab. Kyle Smith for generously offering his encyclopedic knowledge of the literature and for setting aside significant chunks of his time to edit and critique my papers and proposals. Katy Thorn, Terra Barnes, and Hu Dan for tolerating my youthful naivete and patiently teaching me the behavioral electrophysiology techniques. Dan Gibson and Andrew McCool for their programming and analysis assistance (especially for the data presented in Chapter 2) and general knowledge of electrophysiology. Henry Hall for teaching me the ropes in the Graybiel lab and helping me with all manner of miscellaneous (and not always pleasant) tasks from ordering equipment to dealing with animal care issues to building maze contraptions. Yasuo Kubota for painstakingly editing my papers and proposals. Ledia Fernandez for allowing me to collaborate with her on her dopamine depletion studies. Christine Keller-McGandy and other members of the histology lab for performing all the histology on my rat brains. My undergraduate research assistants: Greg Telian, Tshiamo Lechina, and Letitia Li for their hard work and willingness to learn. Alex McWhinnie for general computer and graphic design help. Kostas and Sylvester in the machine shop for their masterful technical work. Greg Hale and Josh Sarinana for helpful presentation feedback and thought provoking scientific discussions (and being great roommates). My committee members, Drs. Matt Wilson, Sebastian Seung, and Ki Ann Goosens for their helpful guidance and penetrating questions. Dr. Lew Haberly, my undergraduate advisor, for inspiring me to pursue academic research and for taking the time to teach me how to do science. My family and my patient fiancé, Lisa, for encouraging me and giving me perspective along the way. And all the Graybiel lab members and members of the MIT-BCS community for being fantastic friends and scientific colleagues.

Chapter 1: Introduction

1.1 Conceptual overview

The brain possesses a multitude of computational resources that allow it to collect information from the external world and transform that information to generate internal perceptions and thoughts. These representations are then used to learn, select, and motivate actions. Brain circuits do not treat all information equally, even at the lowest levels of sensory processing, nor do they produce behavioral output at random. Neural signals are filtered and transformed with a purpose: to produce actions that lead to goals.

Some goals are simple and innate (obtain food, avoid harm, ect.), but many goals are abstract and are learned with experience (getting a Ph.D). The brain constructs predictions about how to achieve these goals based on information from the external environment, and directs and motivates actions based on those predictions. Not all goals are created equal, so the brain utilizes mechanisms for selecting predictions and actions based on their relative values. When the environment changes, or when goals become more or less valuable, the brain is able to adjust its predictions and shift behaviors to adapt to those changes.

How are brain circuits molded to generate the predictions that are needed to direct actions towards goals? One way is through direct experience. When an animal obtains food, for example, feedback signals in the brain could modify synaptic connections so that patterns of neural activity that preceded that food reward are recognized and repeated upon subsequent

trials. This process is termed reinforcement learning (Sutton 1998). Under this framework, repeated experience with goals builds crystallized representations which can be used to guide consistent behaviors. When experience with goals is limited or unpredictable, animals must engage in flexible, exploratory behavior.

In a particular environment, an animal may choose from a variety of alternative actions at any given moment. Exploration means that the course of action chosen by the animal given a particular set of environmental conditions is highly variable. Behavioral exploitation, on the other hand, means that given a set of sensory inputs (state), an animal performs the exact same action or sequence of actions nearly every time. The degree of behavioral flexibility, by this definition, may be directly dependent on the strength of the links between the environmental representations (sensory stimuli, ect.), particular actions, and outcomes (Fig. 1.1). These links may be viewed as "predictions" that vary in a fluid manner over the course of a behavior. The strength of predictions (associative links) may determine the degree to which behavior is exploratory (highly variable) or exploitative (invariant).

The shift from flexible exploration to behavioral consistency or "exploitation" has been widely studied across a variety of animal species (Doupe, Perkel et al. 2005; Graybiel 2005; Graybiel 2008; Humphries, Khamassi et al. 2012). One elegant example is the process by which songbirds learn how to sing. Young birds generate highly variable sequences of song syllables with the goal of matching their vocal output to the song template produced by their parents. They use the auditory feedback from their vocalizations to gradually adjust their song. Eventually, the vocal exploration gives way to a crystallized song structure in adulthood, which

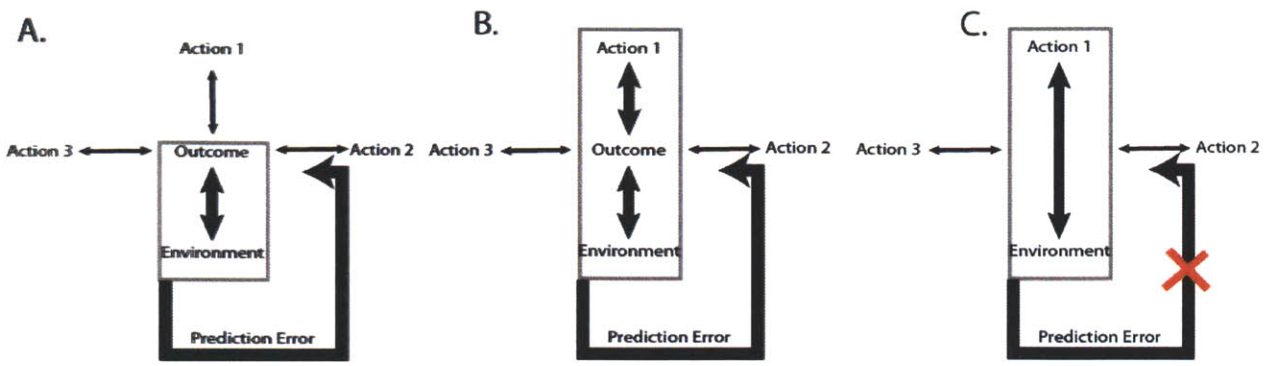

Figure 1.1 Different predictive states that drive behavior. A. In this case, a particular rewarding outcome becomes strongly linked to a predictive sensory state but not to any specific action. The link is built by prediction errors generated by repeated experience with the outcome. Because there are no strong predictive links with any particular action, the animal may engage in variable (exploratory) behavior in this state, aimed at reaching the predictive sensory state. Violations of the predicted association between outcome and sensory state can lead to modification of this prediction via prediction error feedback signals. **B.** If a particular action reliably precedes the unpredicted rewarding outcome, links may be formed between that action and the outcome, in addition to the sensory state as in A. In this case, predictive sensory states may initiate particular actions or action sequences that lead to the predicted outcome. **C.** With repeated predicted exposure to the outcome, action-state links may guide behavior independently of predictions regarding the outcome. Since there is no prediction made regarding the outcome, changes in the reward may not influence the execution of the action, leading to "habitual" behavior.

is relatively inflexible to perturbation. The neural mechanisms involved in this process have been carefully studied and have yielded important insight into how reinforcement learning might occur in mammalian species (Nordeen and Nordeen 1997; Doupe, Solis et al. 2004; Doupe, Perkel et al. 2005; Fee and Scharff 2010).

A slightly different view on behavioral flexibility has come from a classic set of behavioral studies carried out by Dickinson and colleagues, which revealed that prolonged training on instrumental lever press tasks can elicit behavior that is insensitive to changes in the outcome value (Balleine and Dickinson 1998). In this paradigm, rats received injection of lithium chloride, a chemical that induces temporary sickness, after ingesting the reward inside their

home cage. Effects of this "reward devaluation" protocol on lever press behavior were tested at different stages of training and under different schedules of reinforcement on the lever press task. Rats which underwent reward devaluation after prolonged instrumental behavior continued to perform the lever press, while rats that received the reward devaluation after minimal training stopped pressing. Insensitivity to reward devaluation has been used as an operational definition for "habitual behaviors", those that are less sensitive to changes in the outcome and thus inflexible to additional modification (Fig. 1.1). These behaviors are believed to be guided exclusively by links between environmental stimuli and actions.

An important point regarding the above hypothesis is that strong links between an outcome and a particular action or sensory representation do not necessarily mean that the behavior is insensitive to changes in the outcome, as in the reward devaluation studies. In fact, the opposite may be true. When rats were given reward on a variable interval schedule in the instrumental operant task described above (reward delivered at random intervals regardless of lever presses), they were actually less sensitive to outcome devaluation than when they were given reward on a fixed ratio schedule (reward delivered every x number of presses) (Balleine and Dickinson 1992; Balleine and Dickinson 1998). In this case, the link between the outcome and the lever press action was strongest in the fixed ratio case, meaning a stronger outcome "prediction" was present. Changes in the outcome violated this prediction more severely than in the variable interval case, where the prediction (the links between the action/environment and the outcome) was weak (Fig. 1.1 C). These violations of prediction are termed "prediction errors" and are a critical component of reinforcement learning theory (Schultz, Dayan et al. 1997; Waelti, Dickinson et al. 2001; Schultz 2002).

The songbird learning and the instrumental lever pressing studies illustrate important principles that may help to more precisely distinguish flexible from inflexible behaviors. However, most behaviors performed by humans and other mammals are not single fixed action patterns (i.e. a song or a lever press), but rather are fluid movements through space and time that need to be constantly adjusted based on changing environmental conditions. The degree to which a behavior is subject to variation (exploration) may depend on the ongoing strength of the links between the representations of the environment and representations of actions or goals (Fig. 1.1). Stronger links between the environment/actions and outcomes create stronger predictions which drive invariant behaviors but also allow that behavior to be sensitive to changes in the outcome. Not all behaviors can be learned through experience with the desired outcome. Many behaviors are aimed at goals that have never actually been experienced (as in getting a Ph.D). To direct these complex behaviors, the brain must construct abstract models that are composed of multiple predictive layers, a function that is believed to heavily rely on circuits of the neocortex (Miller and Cohen 2001). The evolution and expansion of the neocortex, particularly the prefrontal cortex, has accompanied the emergence of highly abstracted "intelligent" behavior in primates and humans. This thesis will not address the unique predictive capacities enabled by the neocortex. It is very likely, however, that abstract cortical models rely on simple predictive building blocks constructed through the conserved mechanistic principles of reinforcement learning.

1.2 Neural mechanisms

1.2.1 Parallel circuits in the basal ganglia

The neural mechanisms underlying the concepts described in section 1.1 have received increasing attention over the last decade but are still far from a complete level of understanding. There is not one single brain structure responsible for constructing predictions based on sensory experience then translating those predictions into actions – this function depends on the coordinated action of multiple distributed areas. Converging evidence from neurological case studies in humans and brain manipulation and electrophysiological recording studies in rodents have implicated the basal ganglia as being particularly critical for learning and motivating behaviors.

The principle input nucleus of the basal ganglia, the striatum (or caudoputamen in primates) receives converging inputs from a wide range of cortical and subcortical structures, and thus is in an ideal position to integrate many types of information to form predictions or conjunctive associations between stimuli (Shepherd 2004). The projections to the striatum are anatomically segregated: sensory and motor cortical areas project to the dorsal and lateral regions of the striatum (putamen in primates), prefrontal and premotor cortices project to the medial region, and the hippocampal formation, amygdala, and prefrontal cortex project to the ventral regions of the striatum (also called the nucleus accumbens) (Voorn, Vanderschuren et al. 2004) (Fig. 1.2). The outputs of the striatum through the substatia nigra and pallidum to the thalamus remain segregated, creating loops that are believed to process information largely in parallel (Alexander, DeLong et al. 1986).

1.2.2 Segregated behavioral control and information processing in basal ganglia loops

A series of pharmacological inactivation studies in the striatum of rodents and primates have indicated that the segregated cortico-basal ganglia loops may play distinct roles in controlling behavior. Hikosaka and colleagues inactivated the posterior putamen (roughly analogous to the dorsolateral striatum in rodents) and the anterior caudate (analogous to the rodent dorsomedial striatum) in primates trained to perform new and well practiced motor sequences. They observed that manipulation of the posterior putamen disrupted the performance of well practiced sequences while the anterior caudate inactivation disrupted performance of new sequences (Miyachi, Hikosaka et al. 1997). Neuronal recordings in the two regions were consistent with the behavioral findings (Miyachi, Hikosaka et al. 2002). Based on this evidence, they suggested that circuits in the basal ganglia learn procedures of actions in parallel based on different coordinate frames: the anterior caudate uses spatial (or sensory state) information to guide behaviors early in training, while the posterior putamen relies on a motor coordinate frame to guide movement sequences late in training (Hikosaka, Nakahara et al. 1999).

Studies in the rodent have provided further support for the view of parallel processing of information across basal ganglia loops. Inactivation of the dorsolateral region of the striatum results in behavior that is more sensitive to changes in the outcome value, whereas inactivation of the dorsomedial region renders behavior less sensitive to outcome devaluation and also prevents animals from forming associations between actions and particular outcomes during learning (Yin, Knowlton et al. 2004; Yin, Knowlton et al. 2005; Yin, Ostlund et al. 2005; Yin, Knowlton et al. 2006).

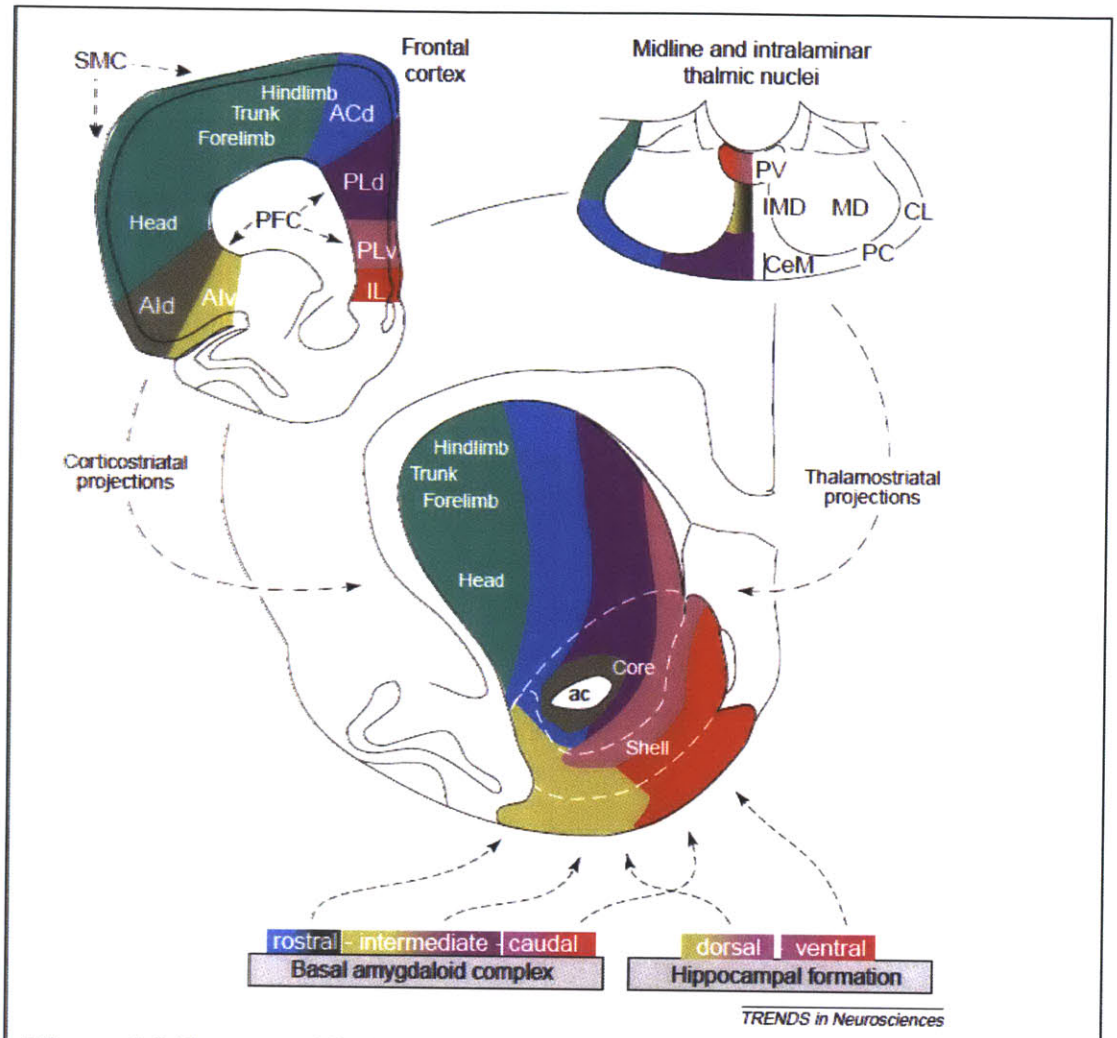

Figure 1.2 Topographical arrangement of projections to the rodent striatum. The schematic shows a coronal cross-section. Afferents are roughly organized along a ventromedial to dorsolateral gradient, with limbic regions projecting mainly to ventromedial and motor and sensory regions projecting to dorsolateral striatum. Adapted from (Voorn, Vanderschuren et al. 2004)

Thus, the dorsolateral striatum may participate in guiding behaviors based on links between stimuli and responses (S-R) without incorporating information about the outcome (i.e. "habits"), while the dorsomedial striatum may participate in behavior guided by links between actions and outcomes (A-O). This is consistent with the idea that parallel basal ganglia circuits utilize different types of information to create predictive associations that can drive behavior.

The ventral medial striatum or nucleus accumbens is notoriously difficult to study using lesion or behavioral pharmacological techniques because these manipulations typically render the animal drastically void of motivation, even to eat and drink. Evidence from electrophysiological studies and specific localized manipulations suggest that this area is involved in learning and representing predictive associations between stimuli and outcomes (S-O) and using these representations to motivate behavior (Salamone and Correa 2002; Nicola, Yun et al. 2004; Atallah, Lopez-Paniagua et al. 2007; Calaminus and Hauber 2007; Humphries and Prescott 2010). The ventral striatum may play a key role in learning instrumental behavior as well (Atallah, Lopez-Paniagua et al. 2007).

Additional evidence for parallel processing of information by striatal learning circuits has come from electrophysiological recordings in behaving animals. In these studies, ensembles of striatal neurons in different subregions were recorded (sometimes simultaneously) for up to several months as rodents learned instrumental T-maze tasks (Barnes, Kubota et al. 2005; Kubota, Liu et al. 2009; Thorn, Atallah et al. 2010). Ensemble firing patterns in the ventromedial, dorsomedial, and dorsolateral striatal regions displayed dramatically different firing patterns and training dependent dynamics as learning proceeded. This data suggests that links between environmental stimuli, actions, and outcomes may be differentially represented across cortico-basal ganglia loops, depending in part on the types of inputs that each area receives. Parallel striatal circuits may also evolve at different rates during learning, depending on the dynamics and the sensory properties of the environment and the actions required to reach the desired goals. Consequently,

behavioral control may be exerted by any of these cortico-basal ganglia circuits, alone or in cooperation.

The strength of the associative links between the sensory and motor representations and the outcome may determine which basal ganglia circuits are most active in controlling behavior. The dramatic effects on motivation produced by the ventral striatum lesions

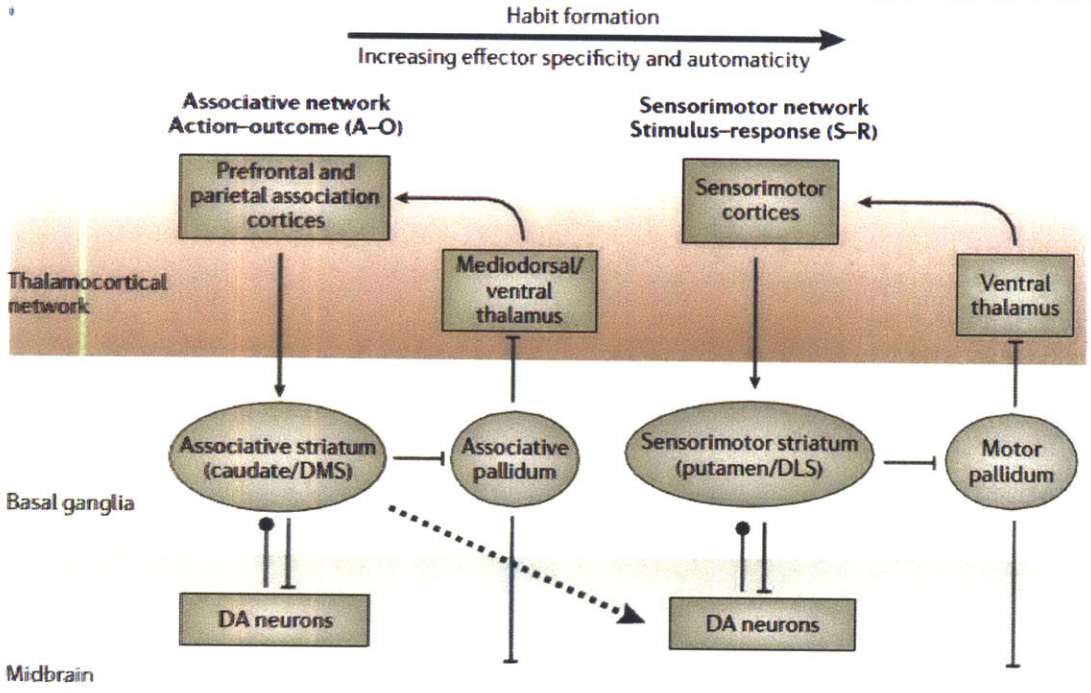

Figure 1.3 Model by Yin and Knowlton for the progression from outcome driven behavior to habitual behavior. Basal ganglia loops involving the dorsomedial striatum (caudate in primates) and prefrontal and premotor cortex form associative links between actions and outcomes, whereas loops involving the dorsolateral striatum (putamen) facilitate primarily sensory state – action links. With increasing training, behavior becomes increasingly insensitive to outcomes (habitual) and is driven more strongly by the dorsolateral circuits. The circuits largely operate in parallel at the level of the basal ganglia but may interact through open loop projections to the dopamine system (dashed line). Adapted from (Yin and Knowlton 2006).

may reflect the crucial role that predictive links between sensory representations and outcomes play in the general invigoration of behavior. Unlike dorsal circuits, which may rely on specific repeatable action representations to be associated with rewarding outcomes, ventral striatal circuits may require only that a particular feature (or combination of features) of the external environment be reliably associated with a rewarding outcome, in many situations, a more common occurrence. Many possible specific action patterns may lead an animal to food, but that food may be associated only (for example) with a specific set of predictive environmental features. This often (but not always) can result in a progression of behavioral control, from ventromedial striatal circuits involved in sensory state dependent control over learning and behavior to dorsolateral circuits involved in action specific, outcome independent control (Hikosaka, Nakahara et al. 1999; Yin and Knowlton 2006; Atallah, Lopez-Paniagua et al. 2007).

1.2.3 A role for dopamine prediction error signaling in reinforcement learning

A critical element of reinforcement learning theory is a signal that carries information about deviations from established predictions, i.e. "reward prediction errors." This signal can be used to flexibly adjust predictions based on feedback from outcome experience. Wolfram Schultz and colleagues carried out a series of landmark electrophysiological studies in the primate midbrain dopaminergic nuclei, the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc), which revealed firing patterns that were highly consistent with the reward prediction error signals postulated by reinforcement learning modeling studies (Fig. 1.3) (Schultz, Dayan et al. 1997; Waelti, Dickinson et al. 2001; Schultz 2002). This work was carried out using a simple classical conditioning task: head-fixed monkeys were given an auditory cue that was followed a

short interval later by delivery of a juice reward. Dopamine neurons were initially fired vigorously in response to the unpredicted juice reward, but with extended training those same neurons began responding to the predictive auditory cue. When the juice reward was withheld suddenly, the dopamine neurons showed a dip in their firing rates at the time the juice was typically delivered. This behavior of the dopamine neurons suggests that they signal positive prediction errors early in training which may be used to drive the construction of predictions about the environment. The firing to the auditory cue may represent a prediction about the upcoming reward (or a prediction error in the timing of the auditory cue). When rewards of different sizes were delivered, both the positive prediction errors to the reward and the learned predictive signals to the cue scaled with the reward magnitude, indicating that predictions and prediction errors can reflect relative values of outcomes (Tobler, Fiorillo et al. 2005). Subsequent work by a variety of labs has shown that these dopamine signals also reflect reward identity, delay to reward delivery, and effort required to obtain the reward (Waelti, Dickinson et al. 2001; Morris, Nevet et al. 2006; Day, Jones et al. 2010; Gan, Walton et al. 2010).

Dopamine neurons project broadly to nearly all cortical and subcortical brain areas (Fig. 1.4), but their projections are strongest to the striatum. There, dopamine is critical for synaptic plasticity via long term potentiation (LTP) and depression (LTD). Stimulation of dopamine neurons to the striatum paired with stimulation of the afferent cortical projections to the striatum induces lasting potentiation of the glutamatergic cortical inputs (Reynolds and Wickens 2002). Altering the timing of the cortical stimulation with respect to the dopamine pulse can switch the effect on synaptic transmission from potentiation to depression (Calabresi, Maj et al. 1992; Calabresi, Picconi et al. 2007; Shen, Flajolet et al. 2008). By these Hebbian mechanisms, dopamine

impulses might strengthen co-occurring inputs to striatal neurons, allowing those inputs to drive responses more effectively on subsequent exposures to the stimuli. The creation of

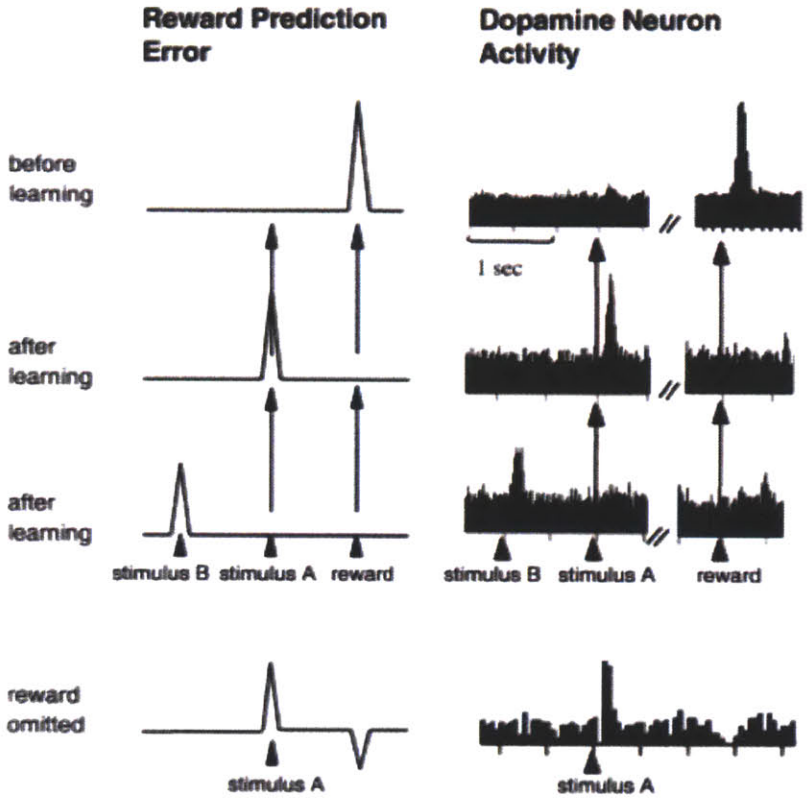

Figure 1.4 Reward prediction error signaling in midbrain dopamine neurons. Electrophysiological recordings were made from putative dopamine neurons in the midbrains of head fixed monkeys performing a classical conditioning task. Histograms in the right column show responses of a representative single unit. Initially, the monkeys were given free unpredicted reward, and the neuron fired phasically at reward delivery (top). When a predictive auditory stimulus preceded the reward delivery at a fixed interval, the neuron fired instead to the stimulus, and no longer to the reward, and when a second stimulus was introduced, the neuron fired to the first stimulus (mid). When the stimulus was delivered but the reward was withheld, the neuron displayed a decrease in firing rate at the usual reward time (bottom). Adapted from (Suri and Schultz 1998)

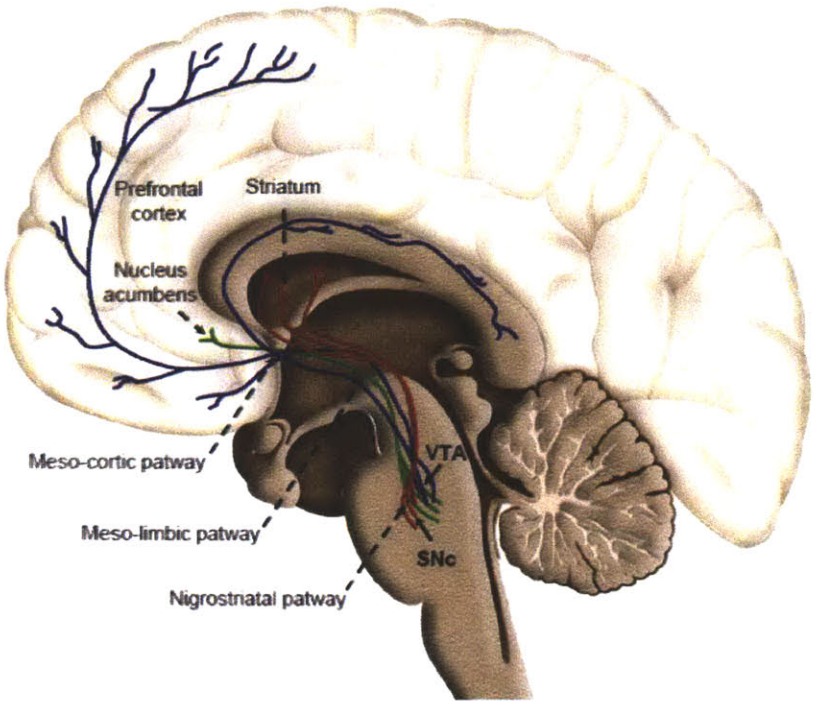

Figure 1.5 Schematic of projections from dopamine systems of the midbrain. Dopaminergic nuclei in the midbrain can be roughly segregated on the basis of their projection patterns. The substantia nigra pars compacta projects mainly to areas of the sensorimotor caudoputamen (striatum) through the nigrostriatal pathway, while the ventral tegmental area projects to limbic and prefrontal cortical structures including the nucleus accumbens, prefrontal cortex, and hippocampus. Image adapted from http://neuro-science.blogspot.com/2011/10/dopamine-and-adult-neurogenesis.html

differential synaptic weight profiles in the striatum may be the neural implementation of reward predictions.

1.2.4 A role for dopamine in motivating behavior

Manipulation of dopamine signaling has provided some support for the role of prediction error signaling in forming reward predictions in the striatum based on external stimuli (Yun, Wakabayashi et al. 2004; Zweifel, Parker et al. 2009; Flagel, Clark et al. 2011; Wang, Li et al. 2011). Mice with specific attenuation of transient dopamine signaling are impaired in acquiring appropriate responses to predictive cues (Zweifel, Parker et al. 2009; Wang, Li et al. 2011). The primary effect of global, non-specific dopamine depletion is not a simple generalized learning deficit, however, but instead is a profound impairment in motivation (Zhou and Palmiter 1995). Pharmacological blockade of dopamine signaling produces impairments in exerting effort to pursue and obtain rewards (Salamone and Correa 2002; Salamone, Correa et al. 2007). Moreover, diseases of the dopamine system, such as Parkinson's disease, produce severe difficulties in initiating and maintaining movements. One possible interpretation of these results is that dopamine is involved in forming but also in *transmitting* predictions about stimuli related to goals. The prediction error signals at goals create predictions (perhaps in striatal circuits) which are themselves transmitted by dopamine neurons in the service of initiating and maintaining behaviors. By this view, dopamine signals to environmental stimuli (such as the signal to the auditory tone in the Schultz studies) are created by reward prediction errors and represent predictions that can directly initiate and maintain actions. A form of this hypothesized role of dopamine signaling has been termed "incentive salience," and has gained increasing support in the literature (Berridge 2007; Flagel, Clark et al. 2011).

A key point regarding the incentive salience hypothesis is that dopamine prediction error and prediction signals need not be necessary for learning or executing (respectively) *specific* action representations (i.e. a sequence of movements) persay. Predictive dopamine impulses to environmental stimuli may have a general invigorating effect on behavior rather than initiating specific behavioral patterns (Schultz, Apicella et al. 1992; Schultz, Apicella et al. 1993; Setlow, Schoenbaum et al. 2003; Nicola, Yun et al. 2004; Berridge 2007). In other words, dopamine may not be required for establishing links between stimuli and responses, but may be important for incorporating reward feedback into predictive sensory representations (Fig. 1.1).

1.2.5 Circuit mechanisms for the influence of dopamine on motivating behavior

Behavioral pharmacology (Calaminus and Hauber 2007; Salamone, Correa et al. 2007; Calaminus and Hauber 2009) and disease states indicate that a main site for the immediate action of predictive dopamine signaling on motivating behaviors may be the striatum itself. One effect of transient dopamine signaling in the striatum is to promote long term synaptic plasticity, as previously discussed, but dopamine signals also have powerful immediate effects on striatal network activity. Signaling through D1 receptors increases and maintains the excitability of medium spiny neurons, while signaling through D2 receptors holds them in a hyperpolarized resting state (Surmeier, Ding et al. 2007; Gerfen and Surmeier 2011). D1 and D2 receptor expression is largely segregated onto different populations of medium spiny neurons (Gerfen, Engber et al. 1990; Le Moine and Bloch 1995). These distinct populations differ also in their output projections to the downstream basal ganglia nuclei. The D1 expressing neurons project to the substatia nigra pars reticulata (SNr) and the internal segment of the globus pallidus (GPi),

while the D2 expressing neurons project to the external segment of the globus pallidus (GPe) (Gerfen, Engber et al. 1990). The net effect of activation of these two pathways on thalamic neurons is believed to be opposite: D1 neuron activation, through the direct pathway, promotes thalamic excitation, while D2 neuron activation, through the indirect pathway produces net thalamic inhibition (Hikosaka and Wurtz 1985; Mink 1996). Moreover, activity in these two pathways is believed to promote opposite effects on movement: the direct pathway promotes movement, while the indirect pathway suppresses movement (Mink 1996; Kravitz, Freeze et al. 2010).

Dopamine signaling, by virtue of its opposing effects on direct and indirect pathway neurons is believed to have a net movement promoting effect.. This hypothesis has gained support from behavioral studies in transgenic mice that express light sensitive channelrhodopsin specifically in either D1 or D2 expressing neurons. Optical stimulation of direct pathway (D1) neurons produces increased movement, while indirect pathway stimulation produces attenuated movement (Kravitz, Freeze et al. 2010). Moreover, a prominent theory of Parkinson's disease states that dopamine degeneration produces hyperactivity in the indirect pathway neurons, resulting in movement suppression (Obeso, Marin et al. 2008). This theory has gained some support from electrophysiological studies in behaving animals which have shown increased firing rates of striatal projection neurons with dopamine depletion (Kish, Palmer et al. 1999).

The feedback interactions between the striatum and dopamine systems are not entirely reciprocal: dopamine projections extend to larger regions of the striatum (and other brain areas) than they receive projections from (i.e. non-reciprocal open loops) (Haber, Fudge et al. 2000;

Joel and Weiner 2000; Matsuda, Furuta et al. 2009). Regional divergence of dopamine projections is present on multiple scales, from inter-regional (on the scale of striosomes and matrix) to cross-regional (ventromedial to dorsolateral divergence). One unique function of the dopamine system then may be to broadcast predictions (and prediction errors) to regions of the striatum which do not themselves encode reward predictions. This would allow the predictive dopamine signals to generally promote ongoing actions and state representations.

Finally, dopamine signaling may be critical for modulating oscillatory network dynamics in striatal networks. Stable oscillatory network states can be created by fine tuned ionic conductances in combination with local network interactions (Wang 2010; Buzsaki and Wang 2012). These oscillatory states have been associated with a variety of functions including memory formation, voluntary movement, and sensory binding (Engel, Fries et al. 2001; Engel and Fries 2010; Wang 2010). A large body of evidence suggests that dopamine signaling may affect the expression of these network states, both positively and negatively (Chapter 2; (Brown 2007; Berke 2009; Benchenane, Peyrache et al. 2010). Thus, the effect of dopamine on the striatal networks may not be as simple as increasing or decreasing firing rates in different populations, but it may fundamentally change the way in which the striatal networks process information.

1.2.6 Reward predictive signaling in striatal circuits

According to models for basal ganglia dependent reinforcement learning, the targets of the dopamine neuron reward prediction error signals should represent reward predictions in their

firing patterns. Electrophysiological recordings of neurons in the striatum in behaving animals have revealed some evidence for predictive representations. Single units in the ventral striatum (primarily the nucleus accumbens core) of both rodents and primates acquire predictive responses to cues during learning of classical and instrumental conditioning tasks and these responses scale with outcome value (Hikosaka, Sakamoto et al. 1989; Schultz, Apicella et al. 1992; Schultz, Apicella et al. 1993; Setlow, Schoenbaum et al. 2003; Nicola, Yun et al. 2004; Goldstein, Barnett et al. 2012). In the dorsomedial striatum, units have been shown to represent contingencies between actions and outcomes (i.e. they fire for particular actions and scale with the outcome value) (Kimchi and Laubach 2009; Kimchi, Torregrossa et al. 2009; Stalnaker, Calhoon et al. 2010; Stalnaker, Calhoon et al. 2012). These observations are consistent with the behavioral work that has implicated the ventral striatum in representing stimulus-outcome contingencies and the dorsomedial striatum in representing action-outcome contingencies (Yin and Knowlton 2006).

The striatum sends projections directly to the midbrain which could, in principle, provide the dopamine neurons with their predictive firing properties. In the dorsal striatum, these neurons are clustered into regions called "striosomes" which are distinguished from the surrounding "matrix" neurons by their differential molecular expression patterns and afferent projections (Graybiel and Ragsdale 1978; Graybiel, Ragsdale et al. 1981; Gerfen 1985). In the ventral striatum, the distinction of striosome and matrix compartments is less clear, but the projections to dopamine nuclei (primarily the VTA) are at least as strong as the dorsal (Kalivas, Churchill et al. 1993; Humphries and Prescott 2010). The projections from the striatum to dopamine nuclei come from GABAergic medium spiny neurons, so direct synaptic connections cannot be directly

translated into the excitatory predictive responses observed in the dopamine neurons (though excitation may be achieved by local interneuron disinhibition). There are a number of indirect routes by which the striatum could communicate reward predictions to dopamine neurons. For example, neurons of the ventral striatum send inhibitory projections to the ventral pallidum, which sends direct inhibitory projections to the ventral tegmental area (Kalivas, Churchill et al. 1993; Humphries and Prescott 2010). Excitatory responses in the ventral striatum would thus generate disinhibition of dopaminergic neurons (Floresco, Todd et al. 2001; Floresco, West et al. 2003). It is likely that multiple convergent pathways from a variety of brain areas, not one particular source, contribute to the predictive responses in dopaminergic neurons (Grace, Floresco et al. 2007).

1.2.7 Brief Summary

-The basal ganglia nuclei, in particular the striatum, contribute both to learning and motivating behaviors towards goals.

-Striatal subregions receive topographically organized projections and this organization is preserved throughout the cortico-basal ganglia circuits.

- Striatal subregions make different contributions to learning and behavioral control and exhibit distinct neuronal firing patterns during behavior.

-The control of different basal ganglia loops over behavior may be determined by the strength of associative links between motor patterns, sensory representations, and outcomes.

-Midbrain dopamine neurons transmit reward prediction error signals, which may be used to establish associative links between reward predictive stimuli, actions, and outcomes through synaptic plasticity in the striatum and elsewhere.

-Reward predictive signals related to environmental stimuli and actions emerge with experience and are represented by neuronal firing in the striatum, particularly in the ventral and medial striatum, and in dopamine neurons (in addition to other brain regions).

-Predictive signals transmitted by dopamine neurons after learning may initiate and motivate behaviors via widespread projections to the striatum and other regions.

1.3 Scientific philosophy and research summary

The objective behind the research described in this thesis is to further the fundamental understanding of the brain mechanisms that guide and motivate human behavior. A comprehensive model would ideally incorporate a minimum set of basic principles needed to achieve general predictive power at the level of neural circuit behavior. Clearly, complex functions like behavioral learning, selection and motivation involve the coordinated action of multiple brain areas. Current methods do not allow us to monitor the neural activity patterns in all of these circuits simultaneously, so we must either study them in isolation (or a few at a time) or construct biologically realistic computational models of neural networks. Both approaches are currently being employed in various forms. One version of the latter relies on a comprehensive database of anatomical connectivity and molecular expression patterns and has been termed "connectomics" (Lichtman and Sanes 2008; Jain, Seung et al. 2010; Denk, Briggman et al. 2012; Reid 2012). The task of collecting this data, cataloging it, and packaging it into informative

computational models is daunting at many levels, but rapid progress is being made on all these fronts. Another similar effort to construct large scale biologically realistic models of the cortex, called the Blue Brain Project, is being carried out by Henry Markram and colleagues (Markram 2012). While computationally enormous in scope, this project does not rely on detailed knowledge of specific synaptic connections like connectomics, but instead uses connectivity and molecular expression patterns of typical cortical microcircuits then constructs a large-scale simulation based on those basic computational units.

A second approach is to measure the electrical and chemical dynamics in neural circuits in awake, behaving animals in regions of the brain known to be critical for guiding behaviors. The idea behind this more focused method is to derive basic principles of circuit operation that might be generalized across multiple behaviors and to large neural ensembles extending across multiple brain regions. These principles would not predict circuit behavior with exact precision but would make general, qualitative predictions. Precise temporal and cell-type specific manipulation of brain circuits, such as with genetically expressed opsins (Boyden, Zhang et al. 2005; Deisseroth, Feng et al. 2006; Witten, Steinberg et al. 2011), would allow us to test the predictions of the models empirically. Such models would be useful not only for understanding some of the basic philosophical questions about human behavior, but would also allow us to understand how circuit operation is disrupted in neurological disorders. It is this general approach that has guided the work presented in this thesis.

In recent decades, a variety of new techniques have been developed that allow us to measure both electrical and chemical dynamics of neural circuits chronically in behaving animals with

high temporal and spatial precision (Robinson, Venton et al. 2003; Buzsaki 2004; Dombeck, Khabbaz et al. 2007). In this thesis, I will describe work using two of these techniques: multi-site tetrode recording and fast-scan cyclic voltammetry to study the electrophysiological and electrochemical dynamics respectively of the striatum in rats during learning and execution of goal-directed behaviors.

The second chapter will describe how oscillations in the local field potential and associated network synchrony in the striatum evolve over the course of learning a simple associative maze task. We have found that early stages of learning are marked by higher frequency oscillatory activity in the gamma range, which gives way to lower frequency beta oscillations over the course of learning. Spiking synchrony shifts during this period as well, from primarily local coordination during the high gamma dominated learning period, to widespread global synchrony during beta oscillations late in training. We propose that a high frequency, local network regime favors plasticity and flexibility early in learning when rewards are unpredicted, but that repeated experience with predicted rewards promotes global network synchrony that favors fixed, inflexible behavior behaviors. I also propose a mechanism by which this transition might take place under the framework of reinforcement learning.

In the third chapter, I will describe the dynamics of dopamine signaling in the striatum of rats performing the same T-maze behavior as in Chapter 1. Dopamine is a neurotransmitter known to be critical for forming predictions and motivating actions to achieve goals. We used chronic fast-scan cyclic voltammetry, an electrochemical technique, to obtain real-time subsecond measurements of dopamine concentration in various regions of the striatum. This study revealed

that dopamine signaling in the maze was not transient in nature, as has been previously described for simpler classical conditioning tasks, but rather increased continuously during maze running and peaked at goal reaching. The signal moreover was sensitive to the maze environment, showing biases for different maze arms that could be manipulated by providing different reward sizes at the ends of the maze. We go on to show, using a combination of electrophysiology and fast scan cyclic voltammetry, that the dopamine signals display an inverse relationship to the beta-band oscillations described in Chapter 1, suggesting that they play a significant role in shaping network behavior. We postulate that the dopamine signals represent predictions about the value of upcoming rewards that vary continuously with the animals' proximity to that reward. Such a signal suggests a mechanism for controlling goal directed behaviors in a flexible, context dependent manner. Finally, I propose a qualitative model by which these dopamine signals could be generated that makes testable predictions about the underlying circuit mechanisms. I hope that this work will stimulate future theoretical and experimental work and contribute to a unifying description of the neural mechanisms underlying behavior.

Chapter 2: Shifts in local field potential oscillations and network synchronization during learning

2.1 Introduction and background

Rhythmic activity in the extracellular local field potential (LFP) is thought to reflect the coordination of local neuronal networks along different spatial and temporal scales. This oscillatory coordination has been proposed to serve as an organizational tool that could, among other things, facilitate long-range communication between brain areas, provide a temporal structure for memory formation, and act as a filter for perceptual attention (Engel, Fries et al. 2001; Buzsaki and Draguhn 2004; Wang 2010). The frequency specific fluctuations in power of LFP oscillations vary widely across brain areas and are highly dependent on the ongoing behavioral state of the animal. Thus, LFP activity in different frequency bands may serve different behavior-specific roles.

In the basal ganglia, oscillations in the LFP have likewise been measured across multiple frequency bands and have been associated with a wide range of behavioral states. Beta oscillations (15-30Hz) in motor cortices upstream of the basal ganglia have been linked with a lack of or suppression of movement initiation (Sanes and Donoghue 1993; Baker, Olivier et al. 1997). In Parkinson's Disease, which is characterized by a degeneration of dopaminergic signaling to the basal ganglia, beta oscillations in basal ganglia nuclei are abnormally strong, indicating that hyper-synchronization of basal ganglia networks in this frequency range may

contribute to the movement difficulties observed in this disease (Brown 2007). However, beta oscillations seem to be important for normal basal ganglia functioning as well, as they have been observed in the striatum, the basal ganglia's principal input nucleus, during sequential movement tasks in primates (Courtemanche, Fujii et al. 2003) and during action selection in rodents (Leventhal, Gage et al. 2012). These studies also indicate that the role of beta oscillations in the basal ganglia is not limited to controlling motor output exclusively.

High gamma oscillations (70-90Hz), common in cortical structures, have also been measured in basal ganglia nuclei, particularly in the striatum. Unlike beta oscillations, high gamma in the striatum has been associated with periods of active movement (Masimore, Schmitzer-Torbert et al. 2005). Moreover, coordination in high gamma between the amygdala and the limbic region of the striatum has been shown to correlate with behavioral performance during Pavlovian association learning (Popescu, Popa et al. 2009), and high gamma power is strongest during the early stages of instrumental learning and wanes with extended training (van der Meer and Redish 2009).

This evidence suggests that different oscillatory network states in the basal ganglia may actively facilitate both network plasticity and behavioral rigidity as a behavior is learned and becomes habitual with overtraining. We tested this hypothesis by recording spiking and local field potentials from the ventral striatum of rats as they learned and eventually habitually performed an associative T-maze task.

2.2 Results

2.2.1 Beta oscillation dynamics during associative T-maze behavior

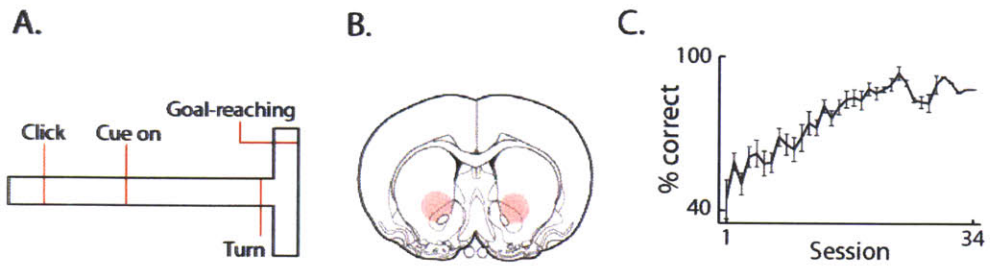

Figure 2.1 Task description and recording sites. A. Layout of the T-maze task. **B.** Region sampled by tetrode recordings (red shading). **C.** Average learning curve for all animals ($n = 7$).

We recorded spike and LFP activity bilaterally from the ventromedial part of the striatum (here called VMS) in 7 rats fitted with chronically implanted headstages carrying assembles of 12 independently movable tetrodes. The rats were trained to perform a T-maze task (Barnes, Kubota et al. 2005) that required them to turn left or right in response to auditory cues in order to receive a reward of either bits of chocolate or a drop of chocolate milk at the end of the arm indicated by the corresponding auditory instruction cue (Fig. 2.1; Methods). The tetrode recordings started at the beginning of training and continued until the rats had completed at least 10 consecutive 40-trial sessions with performance levels above the learning criterion of 72.5% correct (Fig. 2.1*C*). The tips of most tetrodes were confirmed by later histology to be in the VMS, and they were mainly within the core of the nucleus accumbens (Fig. 2.1B, Methods).

Spectral analysis of LFPs recorded after completion of training indicated task-related modulation of oscillatory activity in multiple frequency ranges in the session-averaged LFPs (Fig. 2.2 A and

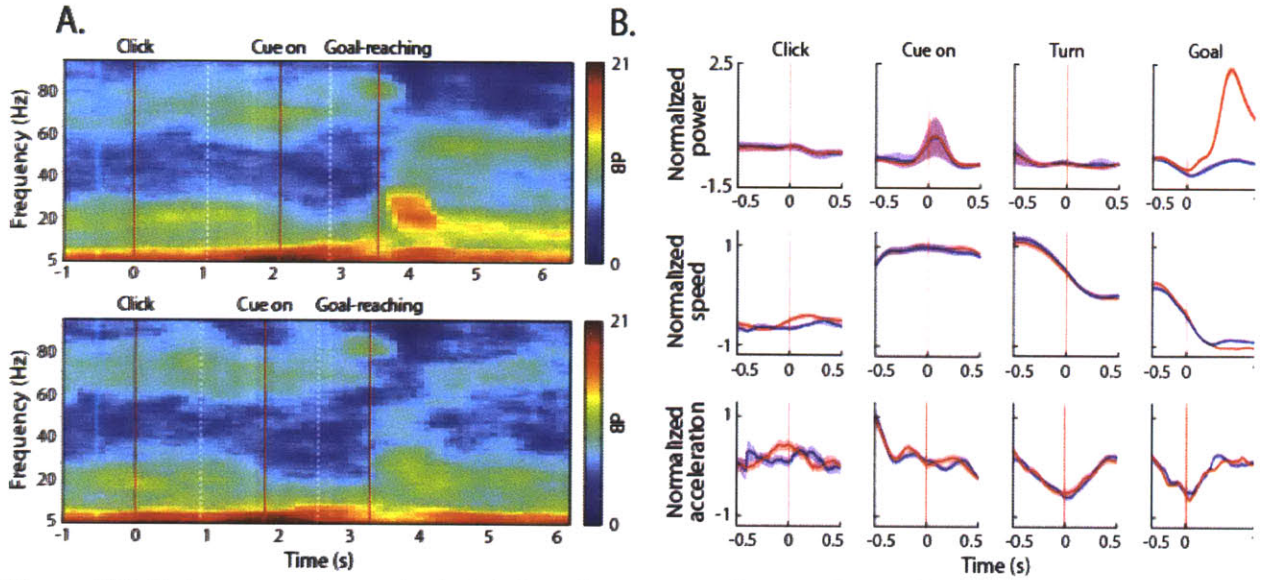

Figure 2.2 Beta power in ventral striatum increases after goal-reaching on correct trials. **A.** Session-averaged spectrograms for correct trials (n = 22 trials, *top*) and for incorrect trials (n = 18 trials, *bottom*) for a representative 40-trial session. **B.** *Top*: Z-score normalized beta power (15-28 Hz) averaged across all sessions and recording tetrodes for correct trials (red) and incorrect trials (blue). Red vertical lines indicate centers of task-event windows as labeled. *Middle*: Average z-score normalized run-speed. *Bottom*: Average z-score normalized acceleration. Shading shows standard errors of the mean (SEM) computed across sessions.

B). Particularly prominent was a transient increase in power in the beta band (15-28 Hz) that occurred directly after goal-reaching (Fig. 2.2A). The increase was strong on correctly performed trials but was weak on incorrect trials (Fig. 2.2 A and B). This modulation of beta-band power did not exhibit a clear relationship to either the run speed or the acceleration of the animals (Fig. 2.2B).

To investigate this beta-band activity in more detail, we examined raw and band-pass filtered LFP traces from single trials. Remarkably, this analysis showed that the extended period of increased beta power visible in the session-averaged spectrograms after goal-reaching did not represent continuous beta-band activity, but instead, corresponded to brief bursts of 2-4 cycles of high-amplitude beta activity (~100-200 ms) that occurred at slightly different times in different trials (Compare Figs. 2.2A and 2.5B).

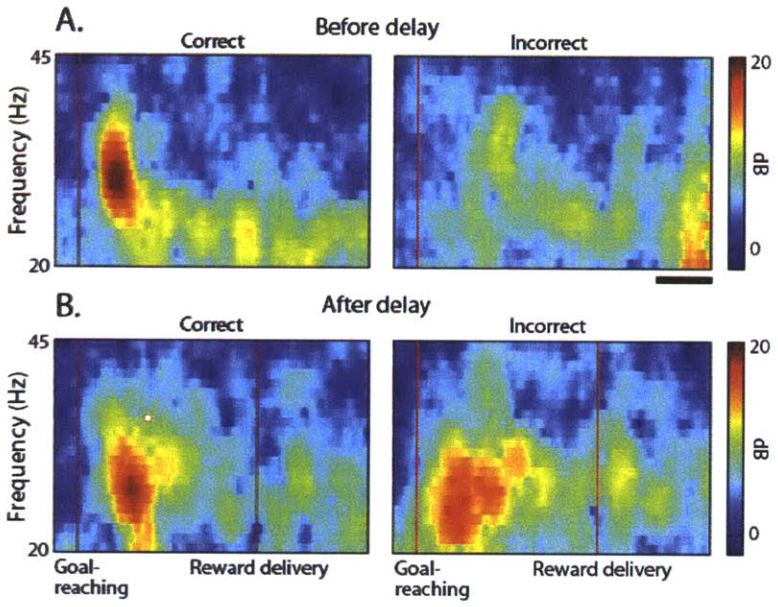

Figure 2.3 Increases in beta power are not dependent on the presence of the primary reward. Reward delay control. Session-averaged beta power for correct (*left*) and incorrect (*right*) trials during a standard 40-trial session (**A**) and during a control session in which reward was delayed for 2 s after goal-reaching (**B**).

To determine whether the transient beta-burst activity was related to reward at the end of the correct runs, we performed two control experiments in which we manipulated the primary reward that the rats received. In the first experiment (n = 2 rats over 13 sessions), we tested whether the sharp increase in beta power was tied to the animals' receiving the chocolate milk reward by delaying the chocolate milk delivery until 2 s after goal-reaching (Fig. 2.3). The increase in beta power after goal reaching did not shift forward in time, but instead, continued to occur just after goal-reaching (Fig. 2.3B). Moreover, beta power in incorrect trials, nearly absent before the delay, became stronger during the trials with reward delay (Fig. 2.3).

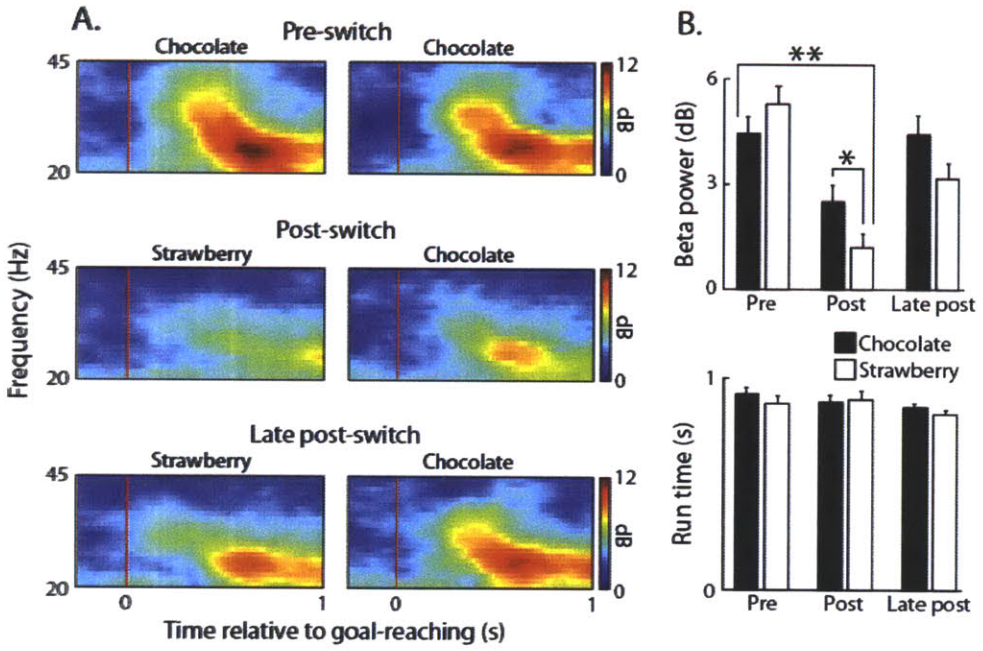

Figure 2.4 Beta power is altered following a switch in reward identity. Reward-switch control. **A.** *Top*: Beta power averaged across correct trials on the left and right sides of the maze (left and right columns, respectively) during two sessions averaged just prior to switching the primary reward. *Middle*: Average beta power after goal reaching for 2 sessions averaged immediately after primary reward on the left arm of the maze was switched from chocolate milk to strawberry milk. *Bottom*: Average beta power for two sessions averaged, 5 days after the switch to the strawberry milk as reward. **B.** Average beta power during the period from 0 to 0.8 s after goal-reaching (*top*) and average run times from turn offset to goal-reaching (*bottom*) for the sessions shown in A. Error bars represent SEM. $*P < 0.01$; $**P < 0.001$.

In the second control experiment (n = 1 rat over 6 sessions), we tested whether changing the identity of the primary reward would influence the beta power after goal-reaching (Fig. 2.4). We did this by suddenly changing the reward in one end-arm of the T-maze from chocolate milk to strawberry milk after the rat had completed ten days of overtraining on the task. Despite the fact that the rat was pre-exposed to strawberry milk in his homecage and drank it readily, we observed a strong and significant ($P < 0.01$, two-tailed t-test) decrease in post-goal beta power on correct trials for both the strawberry milk and the chocolate milk end-arms (Fig. 2.4). This decrease was significantly stronger for the end-arm baited with strawberry milk than for the end-arm baited with chocolate milk ($P < 0.01$, two-tailed t-test; Fig. 2.4B) and could not be accounted for by changes in the run speed of the animal (Fig. 2.4B). After five consecutive days of exposure to the new reward, however, the beta power had rebounded nearly to pre-switch levels (Fig. 2.4). The results from these two control experiments indicate that post-goal beta power is not tied explicitly to the receipt of primary reward but may reflect a particular internal state that can be modified by unexpected changes in the primary reward.

2.2.2 Changes in beta and high gamma oscillations during T-maze learning

We analyzed the entire set of data recorded throughout training to determine whether the patterns of oscillatory LFP activity changed across learning. We found that both beta- and gamma-band trial-averaged activities were composed, when examined trial-by-trial, of transient bursts of oscillations lasting one to several cycles (Fig. 2.5). These bursts happened with varying probability across the T-maze task for the two different frequency bands: gamma bursting was strongest before goal reaching, while beta bursts were more prevalent after goal reaching (Fig.

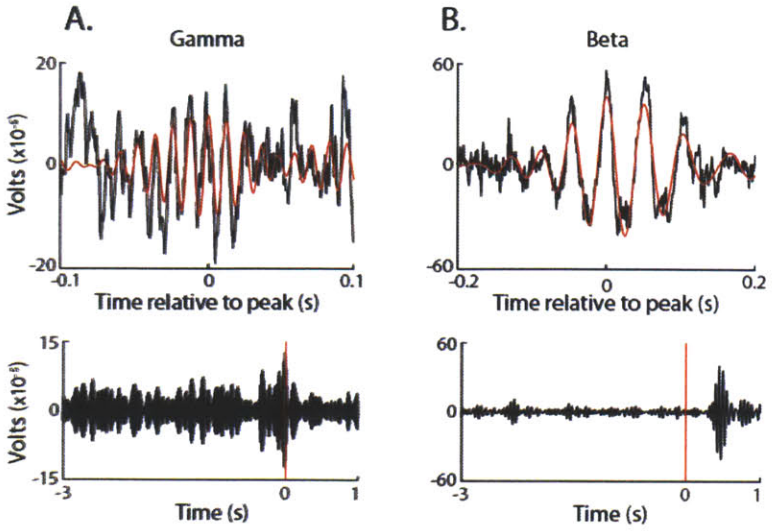

Figure 2.5 Beta and gamma oscillations occur in bursts and are active around different task periods.
Representative examples of single high amplitude gamma (**A**) and beta (**B**) bursts. Top Row: The raw LFP signal is shown in black, and the beta band-pass filtered signal is overlaid in red. Both signals are z-score normalized to the mean and standard deviation of each signal across the entire trial. Bottom Row: Band-pass filtered traces for the high gamma oscillations and beta oscillations. Note the different task modulation for the two frequency bands.

2.6). The occurrence of trial-end beta activity was highly experience-dependent and was accompanied by inversely changing high gamma-band activity (Fig. 2.6). High gamma power was strong both before and after goal-reaching during initial training and the first days of over-training, but it diminished significantly as training progressed. This gamma activity also became more restricted in duration (Fig. 2.6; $R = -0.28$, $P < 0.001$ for before, $R = -0.17$ $P = 0.05$ for after). In sharp contrast, the rise in beta-band power after goal-reaching, typical of correct trials at the end of training, was scarcely detectable even on correct trials at the beginning of training (Fig. 2.6*B*). This beta-band activity did not become prominent during the correct runs until about the time the rats reached the learning criterion, and it then continued to increase throughout the over-training period (Fig. 2.6 *B* and *D*; $R = 0.53$, $P < 0.002$). Beta power during the T-maze runs,

prior to goal-reaching, was much weaker than that after goal-reaching (Figs. 2.6*B*), but it also showed a significant positive correlation with the percent correct performance of the animals (Fig. 2.6*D*; $R = 0.42$, $P < 0.0002$). These findings demonstrate that high gamma power during T-maze performance significantly decreases with learning in the VMS while simultaneously recorded beta power during the T-maze run particularly just after goal-reaching increases with learning.

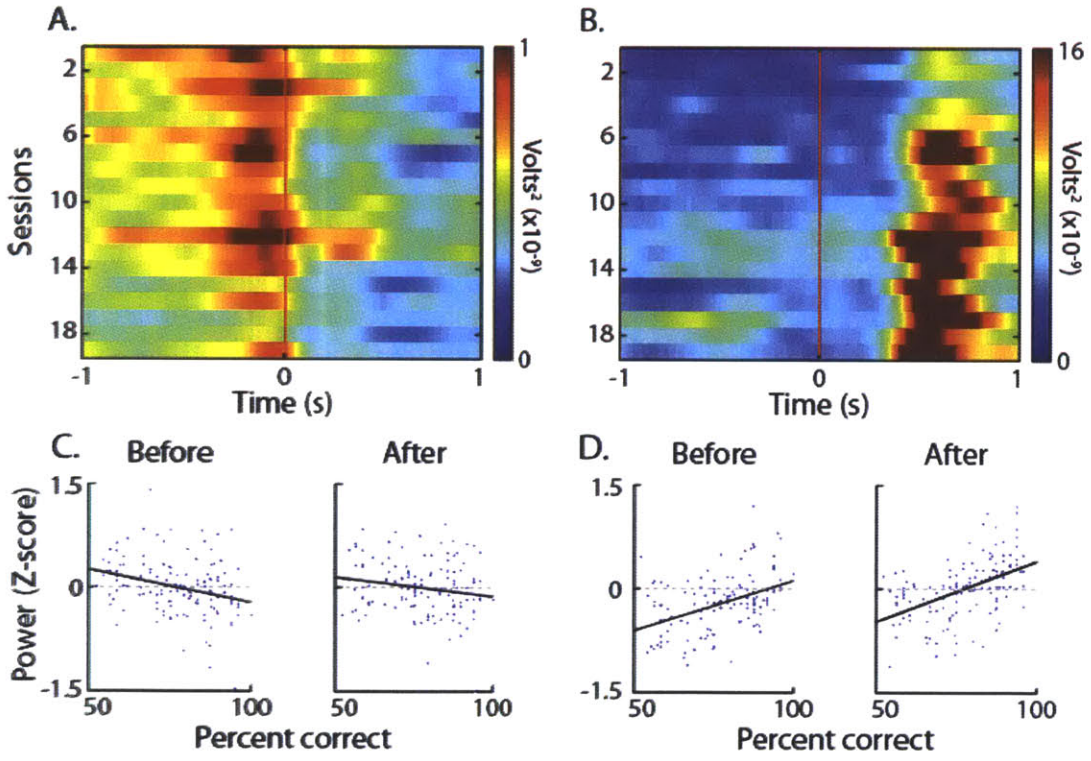

Figure 2.6 Beta power during task performance increases with learning, whereas gamma power decreases with learning. A and B. Gamma (**A**) and beta (**B**) power around goal reaching for all sessions run by one representative rat. Each row represents the average power over 18 correct trials. **C and D.** Mean z-score normalized gamma (**C**) and beta (**D**) power prior to goal-reaching (*left plot*) and after goal-reaching (*right plot*) as a function of percent correct performance for all sessions run by all rats ($n = 7$) combined. Power was z-score normalized in the analyzed window across all sessions for each rat individually.

2.2.3 Putative medium spiny neurons and fast-spiking interneurons are synchronized with beta and gamma oscillations.

We next asked whether the spike activities of striatal fast spiking interneurons (FSIs) and striatal projection neurons recorded were modulated during identified high gamma or beta bursts. We separated neurons into putative subtypes based on their interspike interval (ISI) distributions and their baseline firing rates as described previously (Kubota, Liu et al. 2009) (Methods and Fig. 2.7). Nearly 50% of the FSIs (n = 163/332) were significantly modulated during high amplitude beta bursts, and a similar proportion were modulated around high amplitude gamma bursts (n = 161/332; Rayleigh's test, $P < 0.05$). Just over 50% of the beta-modulated FSIs (98/163) were modulated by both beta and high gamma rhythms, reflecting the overlap in these frequencies mid-learning. Remarkably, as a population, the FSIs with transiently synchronized spiking fired most strongly near the troughs of both high amplitude gamma oscillations and high

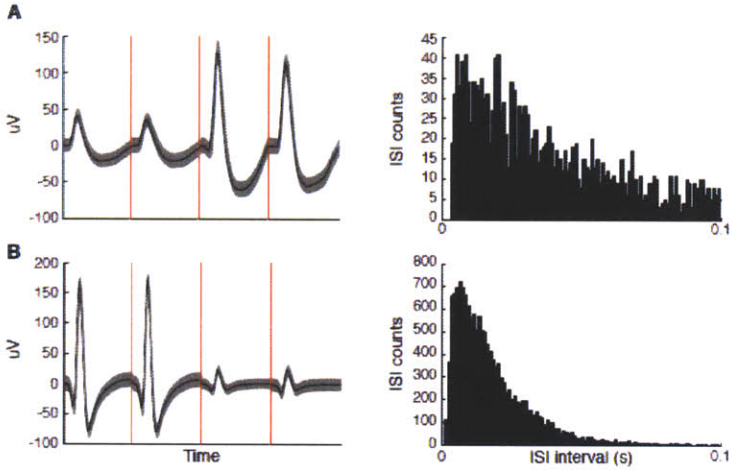

Figure 2.7 Subtypes of striatal neurons were distinguished based on firing rate and interspike interval distribution. Sample waveform (*left*) and distribution of interspike intervals (ISIs, *right*) are shown for a putative medium spiny neuron (***A***) and a putative fast spiking interneuron (***B***). Note the smaller percentage of long ISIs for the putative fast spiking interneuron relative to those for the putative projection neuron.

amplitude beta oscillations (Fig. 2.8 A and B). Given the transition in oscillatory activity in the LFPs during learning, these findings indicate that as a population, the FSIs tended to fire near the phase troughs of high gamma oscillations early in learning but tended to fire at the troughs of beta oscillations late in learning.

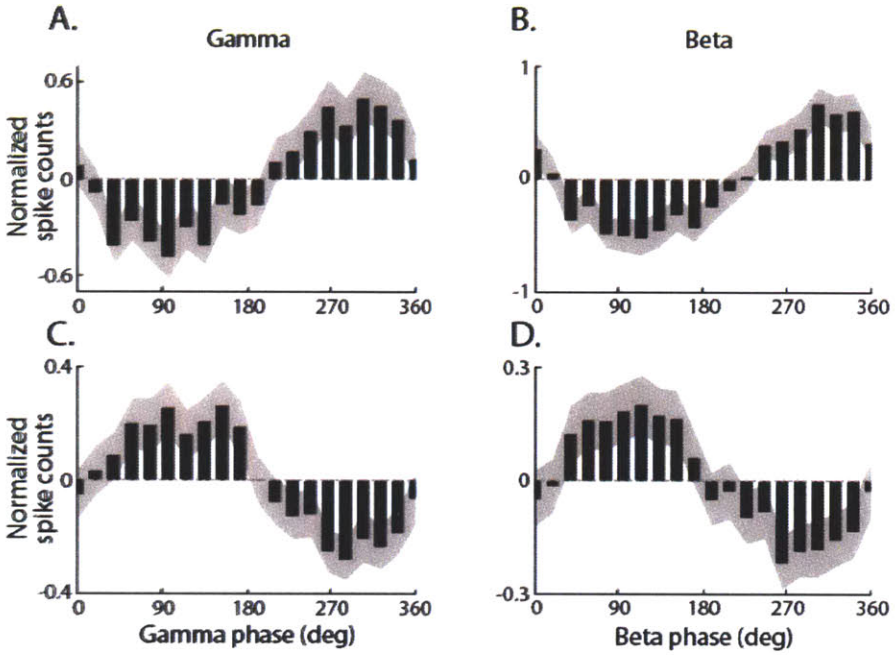

Figure 2.8 Spikes of putative fast spiking interneurons and medium spiny neurons are inversely but coordinately modulated during high amplitude beta and gamma oscillations. A. Mean z-score normalized spike-phase histograms averaged for all significantly modulated FSIs (n = 161) around periods of high amplitude gamma oscillations. Zero phase is defined as the positive-going zero crossing, so spike-phase distributions of FSIs are anti-phase to the LFP oscillation. Z-scores for the spiking of each FSI unit were calculated relative to the mean and standard deviation averaged across spike counts for all phase-bins. **B.** Spike-phase histograms constructed as those in A for significantly beta-modulated FSIs (n = 163) around periods of high amplitude gamma oscillations. **C.** Spike phase-histograms for all gamma-modulated projection neurons (n = 488) around periods of high amplitude beta oscillations. **D.** Spike phase histograms as in *C* for all beta-modulated projection neurons (n = 783). Shaded lines in *A-D* represent 2 x SEM.

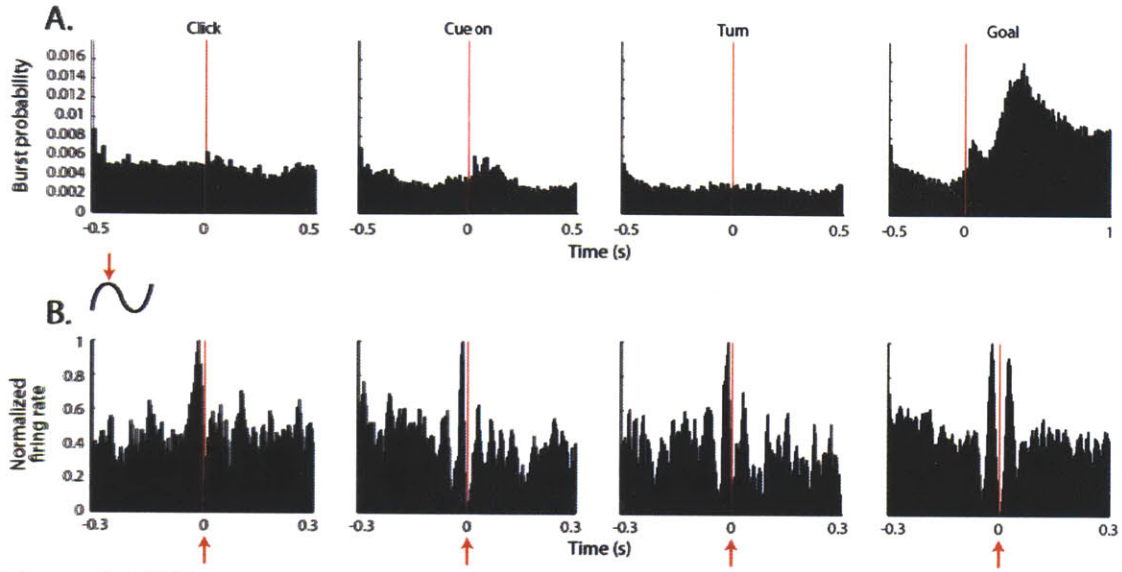

Figure 2.9 FSIs are synchronized during high amplitude beta bursts around each task event analyzed. A. Beta-burst probability during 10 ms-wide bins centered around successive task events in the maze task trial and session averaged across all rats. Bursts were identified by periods during which the z-score normalized band-pass filtered trace exceeded 2.5 standard deviations (SD) above the mean for the trial. **B.** Averaged z-score normalized spike histograms aligned to peaks of beta bursts around each task-event as shown in *A*, for all beta-modulated FSIs (n = 163) across all rats and sessions combined. Modulated FSIs fire around the troughs of the beta bursts occurring around each task-event examined.

In the dorsal striatum, neurons exhibiting at high firing rates in slice preparations have been identified as interneurons that have functional monosynaptic inhibitory connections with the projection neurons of the striatum (Kita, Kita et al. 1985). Despite this finding *in vitro*, few direct inhibitory interactions between projection neurons and FSIs have been reported *in vivo* (Kalenscher, Lansink et al. 2010). We reasoned that if the FSIs that we recorded also inhibited projection neurons, then the spikes of the projection neurons might also be synchronized to the beta burst oscillations, but in the opposite phase to the FSI spikes. We tested this idea and found it to be correct for many of the projection neurons. As a population, over a quarter of the projection neurons showed spike-field modulation (about 28%, n = 783/2796

units) during beta bursts and about 17% (488/2796) during gamma bursts, and the phase distribution of their spikes was opposite to that of modulated FSI spiking during beta bursts: the projection neurons fired near the positive peaks in the LFP (Fig. 2.8 *C* and *D*), regardless of the oscillatory frequency. This anti-phase relationship between FSIs and projection neurons may reflect intermittent inhibition of projection neurons by FSIs. Importantly, modulation was observed not only for periods in which the beta oscillations were most prominent (after goal reaching) but during all task events to some degree (Fig 2.9). This indicates that the learning related effects on beta and gamma oscillations may represent general "state" related changes rather than simply shifts in responsiveness to different events in the task.

2.2.4 Beta bursting and gamma bursting differ in their spatial synchrony across the striatum

We analyzed the spatial distribution of synchronization patterns of the beta and gamma bursts recorded on different tetrodes to estimate the spatial extent of the rhythmic oscillations in the VMS. Beta-band oscillations were highly synchronous across the recording sites (Fig. 2.10). Synchrony was much stronger during high amplitude beta bursts than during random periods outside of beta bursts (Fig. 2.10*B*). By contrast, during high amplitude gamma bursts, synchrony across tetrodes spaced through the VMS was comparatively weak (Fig. 2.10 *C* and *D*). Thus, the gamma activity that appeared early during behavioral learning appeared to be local, whereas the beta activity that appeared later during learning appeared to be more widely distributed across the VMS.

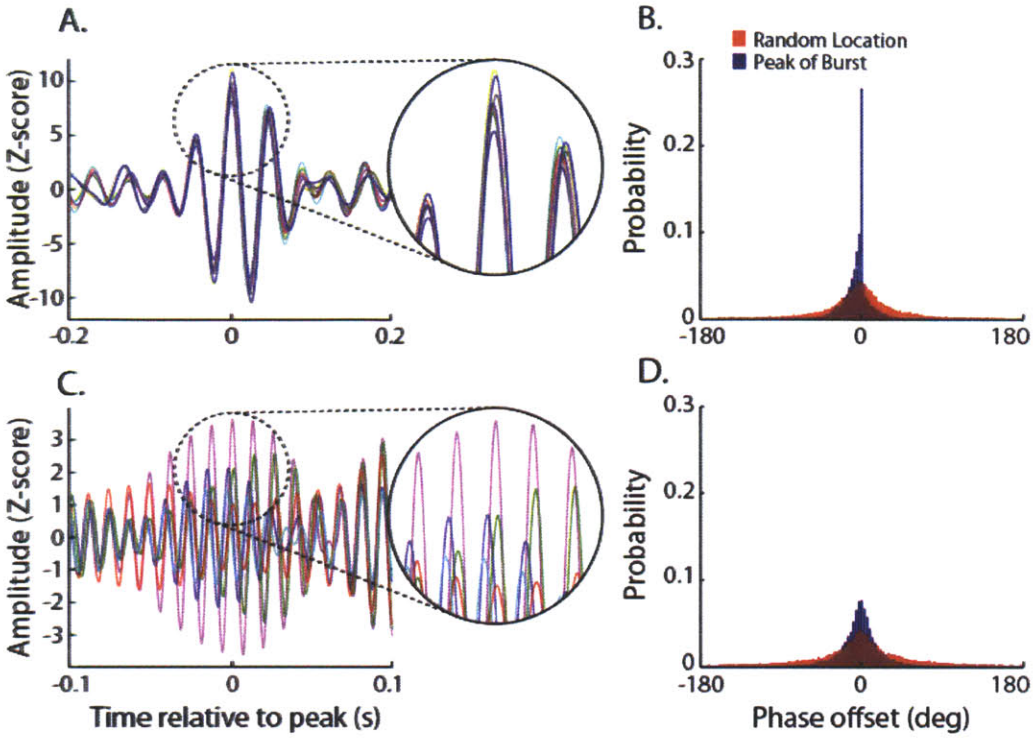

Figure 2.10 Bursts of high amplitude beta oscillations in the LFP are globally synchronized, whereas bursts of high gamma oscillations are more spatially localized. A. Overlaid beta band-pass filtered traces from 8 simultaneously recorded LFPs recorded on spatially separated tetrodes in the VMS around a high amplitude beta burst in a single trial. Traces are centered on the peak of the highest amplitude signal and are z-score normalized to the filtered trace for the entire trial. Insets to right show tight alignment of the peaks on different channels. All LFPs were referenced to an external ground on the recording system. **B.** Histogram of phase offsets on all LFP channels relative to the peak of the highest amplitude LFP trace (reference phase = 90°) during all identified beta bursts (> 2.5 SD above mean, blue) and during an equal number of randomly selected non-burst periods during the task (red). Phase differences around 0 indicate strong synchrony across electrodes. **C and D.** Equivalent plots as *A* and *B* for gamma oscillations, to illustrate looser alignment and lack of heightened synchrony during bursts relative to beta. This indicates that gamma bursts are more spatially heterogeneous than the globally synchronized beta bursts.

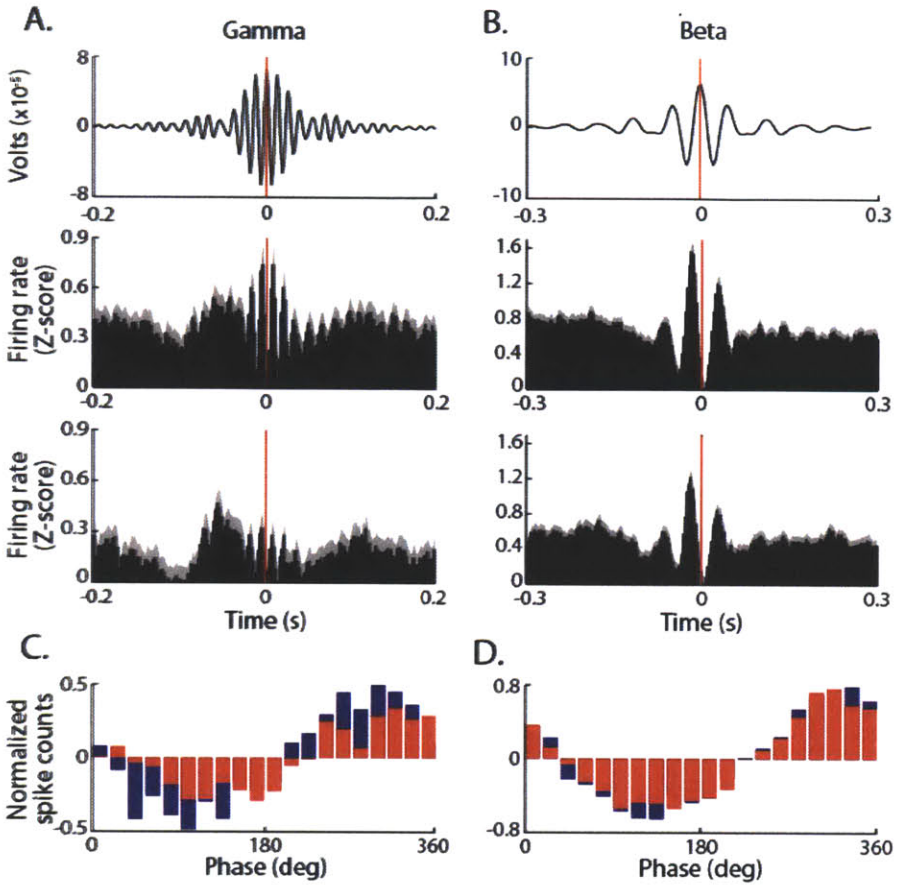

Figure 2.11 FSIs synchronize globally during high amplitude beta bursts, but are more weakly spatially synchronized during gamma bursts A. *Top*: Average band-pass filtered gamma trace used to construct histograms below. *Middle*: Normalized spike histograms of all gamma-modulated FSIs aligned on the peaks of all high amplitude (> 2.5 SD above trial mean) gamma bursts on the local tetrode on which the spikes were recorded. *Bottom*: Normalized spike histogram as above aligned to peaks of high amplitude gamma bursts recorded on a different randomly selected tetrode from the one the spikes were recorded. Baseline subtracted firing rate means calculated around the peak of the oscillation were significantly different between local and non-local alignment ($P < 0.0001$, Kruskal-Wallis test), indicating an absence of global synchrony of FSIs during gamma bursts. Shaded bars represent SEM. **B.** Normalized firing rate histograms, as in A, for high amplitude beta bursts. Baseline subtracted firing rate means calculated around the peak of the oscillation were not significantly different between local and non-local alignment (P =0.48, Kruskal-Wallis test) indicating global synchrony of FSIs during high amplitude beta bursts. **C and D.** Spike-phase histograms, as in Fig. 4, constructed around gamma (*C*) and beta (*D*) bursts recorded on the same electrode as the spiking (blue) and randomly selected tetrodes spatially separated from the tetrode the unit was recorded on during local burst times (red).

We asked whether this spatial disparity in synchrony between beta and gamma bursting also held for spiking. We aligned the spikes of FSIs to the peaks of high amplitude beta and similarly to the peaks of high gamma oscillations (Fig. 2.11 A and B). Modulated FSIs displayed a robust oscillation when aligned to peaks of both beta and gamma bursts on the same tetrode from which the spikes were recorded. When spikes were aligned to peaks of gamma bursts recorded on other spatially separated tetrodes, however, the synchronous oscillation was significantly reduced around the oscillation peaks (Fig. 2.11A; $P < 0.00001$, Kruskal-Wallis test on baseline-subtracted means around peak). By contrast, the spike-field synchrony remained robust when we aligned the spikes of beta-modulated units to peaks of beta bursts on non-local tetrodes (Fig. 2.11B; $P = 0.48$, Kruskal-Wallis test). These findings may help to account for the variable detection of FSI synchrony (Berke 2008; Gage, Stoetzner et al. 2010): the global synchrony of FSIs that would facilitate this detection is characteristic of transient, 100-200 ms periods in which beta bursts occur.

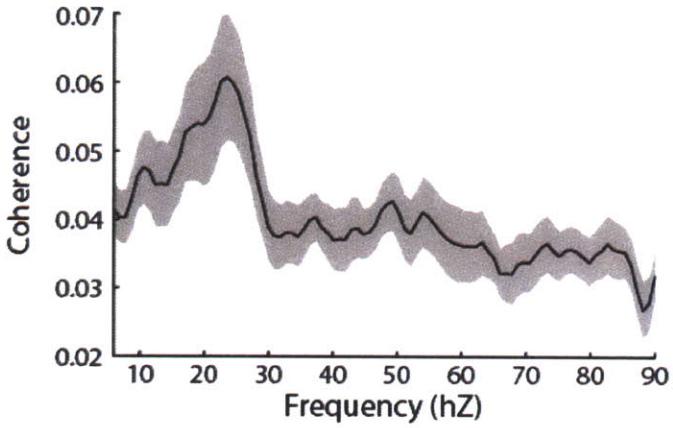

Figure 2.12 Pairs of simultaneously recorded FSIs are coherent at beta frequency. Average spike-spike coherence for all pairs of beta-modulated FSIs that were recorded on different tetrodes simultaneously in single training sessions (n=21 pairs). Coherence was computed for periods of +/- 300ms around peaks of high amplitude beta bursts. Error bars represent SEM.

To quantify this spatial effect on spike-field synchronization further, we constructed spike-phase

histograms as illustrated in Fig 2.11, but we plotted the spikes relative to non-local band-pass filtered LFPs during the high amplitude beta and gamma events (Fig. 2.11*C* and *D*). We found that non-local spike-phase relationships during high amplitude gamma oscillations were significantly weaker than local spike-phase relationships (Fig. 2.11*C*; $P < 0.01$, two-tailed t-test on the absolute values of mean phase distributions). However, such reductions did not occur for the spikes synchronized to the high amplitude beta oscillations recorded (Fig. 2.11*D*; $P = 0.57$, two tailed t-test). These spikes showed a nearly equivalent magnitude of spike-phase modulation for local LFPs and non-local LFPs. Finally, to test whether FSIs are coordinated with each other particularly during beta bursts and to rule out any potential confounds with volume conducted LFPs, we computed spike-spike coherence for pairs of FSIs (Fig. 2.12). Consistent with our previous results, spike-spike coherence peaked around the beta frequency, indicating that FSIs synchronize preferentially in the beta frequency range. These findings suggest that the brief bursts of beta-band oscillation represent periods of global synchrony in the VMS network, whereas high amplitude gamma oscillations primarily represent local synchrony.

2.3 Discussion

Our findings demonstrate that as animals learn a simple T-maze task, the temporal and spatial organization of LFP oscillations in the ventromedial striatum undergoes a major transition, whereby task-related bursts of oscillatory activity in the high gamma range are supplanted by beta-band bursts near the completion of successful runs of the maze task. Populations of striatal interneurons and projection neurons synchronize to opposite phases of both of these rhythms, with the consequence that over the course of learning, spike-field synchrony shifts frequency

bands as well. As our evidence favors spatial heterogeneity of gamma burst synchronies early in learning, and more distributed and uniform beta burst synchronies late in learning, our findings suggest that both the spatial and temporal structure of neural processing profoundly changes during learning.

2.3.1 Beta band oscillations occur in brief bursts, marked by transient population synchrony

Our findings establish that beta-band oscillations in the ventromedial striatum occur predominantly in brief, ca. 100-200 ms bursts. Remarkably, we found no evidence for prolonged periods of beta oscillations when we analyzed the data trial by trial, a finding matching the observations in macaque monkey made in our laboratory and others (Dejean, Arbuthnott et al. 2011). The transient beta bursts that we observed directly corresponded to transient episodes in which the spikes of striatal fast-firing interneurons and projection neurons were synchronized to the beta-band oscillations. Spiking of the fast-spiking interneurons was concentrated near the troughs of the beta bursts, whereas spiking of the projection neurons was strongest near the peaks. Because fast spiking interneurons are known to make monosynaptic inhibitory connections with projection neurons, but no equivalent functional connection has been found from projection neurons to interneurons (Chuhma, Tanaka et al. 2011), it is likely that the alternating periods of firing at the beta frequency are generated by an interneuron-driven suppression of projection neurons. Thus, the LFP beta bursts, lasting one to a few cycles of 15-28 Hz oscillation, likely correspond to fleeting episodes in which the main output neurons of the

ventromedial striatum are rhythmically and synchronously suppressed while populations of inhibitory fast-firing interneurons are synchronously excited.

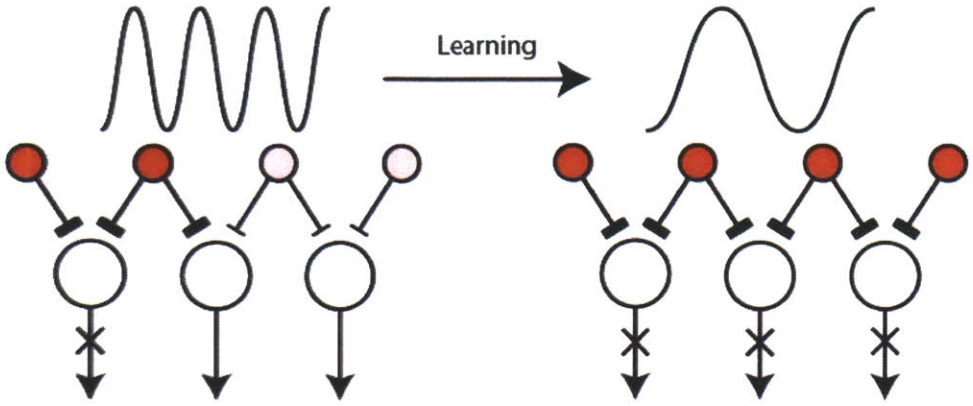

Figure 2.13 Model for changes in spatiotemporal dynamics of oscillatory LFP activity and spiking in ventromedial striatum during habit learning. The left column represents activity patterns early in learning, when high amplitude gamma bursts in LFPs recorded in the ventromedial striatum are prominent at task-end. The right column represents LFP activity patterns after over-training and habit formation, when beta bursts have become strong at task end. Large open circles represent projection neurons, and smaller shaded circles represent inhibitory interneurons. The intensity of shading represents the probability of firing at a given time point during a high amplitude gamma (*left*) or beta (*right*) event. A projection neuron is inhibited if it receives strong (dark red, dark projection lines) inhibition from 2 converging interneurons. Early in training, during high gamma bursts, interneurons synchronize with each other only locally in small groups, and thus provide only local, non-convergent inhibition to small populations of projection neurons. Late in training, during high amplitude beta bursts, interneurons synchronize their firing across wider spatial areas and across a longer cycle period, and as a consequence, project powerful convergent inhibition to a larger population of projection neurons. The result of this beta range synchrony is a widespread inhibition of ventral striatal output during habit formation.

The synchronization of striatal interneurons and projection neurons associated with beta bursts was experience-dependent. The bursts of beta-band LFP oscillation developed as behavioral learning of the task advanced. Moreover, the beta bursts occurred with highest probability and amplitude at task-end, and occurred only in correctly performed rewarded trials. Reward

likelihood was thus clearly critical for the post-goal beta bursts. The nature of its effects was not simple, however. The presence and identity of the reward modulated the strength of the bursts, but the reward-delay trials indicated that the beta bursts were not tightly time-locked to reward delivery itself, and they were also not time-locked to the licking or consumption of the reward. It seems more likely that the bursts, and accompanying synchronizations of spiking, mark expected task-end when the animals are in a particular, habitual mode of successful performance.

These characteristics suggest that the prominent beta bursts that occur at task-end represent brief neural events in which striatal networks can be reconfigured and reset to reflect successful completion of a learnt behavior. This idea could be seen as fitting the proposal that beta-band oscillations post-behavior serving a function in maintaining the status quo (Engel and Fries 2010). Beta bursts and associated spike synchronization did occur at other task-times, but the timing of these bursts was more variable than that of the beta bursts occurring at task-end after learning (Fig. 2.9). This suggests that the beta bursts are not present only during particular times in the task or in response to specific stimuli, but may reflect a particular global network state. We note that it could be critical to the function of these beta bursts that they are associated with transiently and coordinately synchronization of different sets of striatal network neurons: one set, the FSIs, synchronizing near the troughs of the beta oscillations, and another set, the output neurons, synchronizing near their peaks.

2.3.2 Gamma-band activity and beta-band activity are inversely modulated through the course of learning.

Our findings confirm those of van der Meer and Redish (van der Meer and Redish 2009; van der Meer, Kalenscher et al. 2010) that high-gamma oscillations are dynamically modulated during maze learning, decreasing as learning proceeds. By making detailed comparisons of the high gamma bursts and beta-band bursts that we recorded around goal-reaching, we found that there was an inverse relation between the two over the course of learning. Gamma power was strongest early in learning and became weaker during overtraining, whereas beta power was relatively weak early on but became progressively stronger with training. Remarkably, we found that the spikes of the FSIs and projection neurons had the same inverse phase relations to the gamma-band oscillations as they had to the beta-band oscillations: the FSIs spiked near the phase troughs of gamma oscillations, and the projection neurons near the phase peaks. This parallel linkage of the neuronal spiking to the gamma and beta oscillations suggests that there was a major switch during learning from spike-timing regulation by bursts of high gamma early in training to regulation of spike-timing by beta oscillations later in training, after behavioral accuracy had reached asymptote. It is possible that the major consequence of the shift in frequencies was to alter the network from a state of heterogeneous, spatially restricted operation (the gamma regime) to a state of widespread, spatially homogneous processing. It is notable that the changes in oscillatory activity were not limited to the times in which the oscillatory power was strongest but rather were observed at all periods during the task. This suggests that the change that occurs through learning may represent a global state change, rather than task-specific changes. There was a prolonged period during overtraining when these oscillatory activities overlapped. Nevertheless, our observations based on analysis of the spikes occurring during high-amplitude bursts of gamma and beta strongly suggest that spike-LFP coordination itself undergoes a major reorganization during learning.

We did not identify the type of fast-firing units that we recorded other than by their rates of firing and waveforms as recorded by the tetrodes. High and low gamma activities in the nucleus accumbens have been associated with different types of interneurons by van der Meer and Redish (van der Meer and Redish 2009). Thus it is possible that the shift in oscillatory control of spike timing that we observed was accompanied by a shift in the types of FSIs aligned to the gamma and beta oscillations. However, as many FSIs aligned their spikes to both rhythms, there cannot be a one-to-one correspondence between those two FSI types and rhythm preference.

2.3.3 Changes in beta and high gamma oscillations during T-maze learning

We found a striking difference in the spatial patterns of synchrony for the LFP oscillations recorded in the high gamma and beta-band ranges. The early-appearing gamma rhythms appeared to synchronize mainly locally, in agreement with observations by Kalenscher et al. (Kalenscher, Lansink et al. 2010), whereas the late-appearing beta oscillations were broadly synchronized across tetrode recording sites spread across the ventromedial striatum. This difference in spatial structure of the synchrony patterns could have major implications for the accompanying neural processing. As shown in the schematic model of Fig. 2.13, the spatiotemporal patterning suggested by our findings is one in which transient local gamma-synchronized spike-field patterns are strongly present during early learning, whereas transient bouts of beta-synchronized spike-field synchrony, most prominent at task-end, are widely coordinated across the ventromedial striatum late in learning.

What could be the mechanistic significance of this transition from frequent local gamma synchronization to more global beta-synchronization of spike firing? One possibility, raised by numerous previous studies (Engel, Fries et al. 2001; Buzsaki and Draguhn 2004; DeCoteau, Thorn et al. 2007; Popescu, Popa et al. 2009; Wang 2010) is that oscillatory activities facilitate, in frequency-dependent ways, communication between interconnected brain regions. By this view, the shift from high gamma to beta burst synchronization could reflect a shift in the functional connectivity of ventral striatal networks during learning.

Another possibility is that the oscillations represent a particular state of microcircuit operation within the striatal network itself. The shift in balance between gamma and beta during learning appears to reflect a shift from a spatially asynchronous mode of transient FSI firing to a mode in which FSIs transiently become globally synchronized. By this view, early in the exploratory phase of learning, when gamma is strong, different subsets of projection neurons might independently communicate with local interneurons. This type of network environment could favor plasticity and flexibility (Wespatat, Tennigkeit et al. 2004; Headley and Weinberger 2011). When, later in learning, beta oscillations emerge, they might promote a more homogenous network structure and a crystallized, more nearly fixed pattern of output as behavior becomes habitual. This type of network structure may be ideal for promoting stability.

A third possibility is that the widespread FSI synchrony during beta bursts, by allowing a longer time for spatial and temporal summation of FSI spikes than that during gamma bursts and thus a larger pool of interneurons firing together, could promote more effective inhibition of projection neurons through enhanced spatial and temporal summation of the inhibitory post-synaptic

currents (Fig. 2.13). FSIs, which synchronize strongly as a population during beta bursts, fire on average at around the beta frequency (Berke 2008; Kubota, Liu et al. 2009; van der Meer and Redish 2009). A powerful common input to a large population of FSIs, as has been found for some cortico-striatal projections (Parthasarathy and Graybiel 1997), could thus produce a global resetting of FSI firing and transiently synchronize spiking of the entire population at around their intrinsic mean firing frequency (Fig 2.14). Thus, as learning proceeds, the output in the ventral striatum could become progressively dampened.

2.3.4 Relationship of the oscillatory network states to reinforcement learning theory

Much evidence suggests that as habits and procedures are acquired, activity in the ventromedial striatal region, including the nucleus accumbens core, is essential early in the learning process, but that this activity may not be required late in learning (Hernandez, Sadeghian et al. 2002; Atallah, Lopez-Paniagua et al. 2007). In work described elsewhere, we demonstrate that as learning proceeds, the projection neurons in the ventromedial striatum at first fire in relation to goal-reaching, but that this response does indeed nearly disappear as the animals are over-trained. This goal reaching activity appears to relate to reward feedback (i.e. is only present on rewarded trials), which may be used in the service of reinforcement learning as a reward prediction error similar to that observed in dopamine neurons of the midbrain (Schultz 2002). Like the dopamine neuron activity, the neural responses in the VMS diminish as the reward becomes predictable, suggesting that information about the presence or absence of the reward late in training may no longer be immediately recognized by the VMS learning circuits. Insensitivity to changes in reward outcomes is a hallmark of habitual behavior, in which highly

overtrained behaviors continue to be performed in spite of drastic devaluation in the reward associated with those behaviors (termed "reward devaluation") (Balleine and Dickinson 1992; Yin and Knowlton 2006).

The result here suggests a mechanism by which repeated exposure to rewards may induce an "anti-exploratory" network state. In this environment, rewards are fully predicted, so plasticity and network variability (promoted by the gamma dominated network state) would be detrimental to goal-seeking. Repeated network plasticity might favor global network reset by gradually strengthening different subsets of reward activated inputs in an iterative process. Initially reward related inputs are widespread and sparse, favoring sparse MSN excitation (reward related feedback that can propagate out of the network) and network output rather than significant FSI mediated inhibition. Over time, however, synchronous input to large areas of the striatum becomes dominant, favoring synchronous reset of FSI populations, frequent beta burst events and powerful convergent inhibition.

Finally, one consequence of this process is that activity related to the reward may be progressively dampened as repeated rounds of plasticity shift the network towards global FSI synchrony. It's possible that sensory inputs related to the rewarding experience initially drive ventromedial MSN output early in learning, which then induces dopamine release downstream. The same inputs after repeated dopamine induced plasticity (many experiences with the initially unpredicted reward), may increasingly drive synchronous FSI activity which has the effect of canceling the MSN output and reward related dopamine release. The result is a pattern of dopamine activity that recapitulates the reward prediction error firing described in Chapter 1.

2.3.5 Relationship of the beta oscillations observed here to beta observed in Parkinson's disease.

Beta oscillations are abnormally strong in the basal ganglia of patients with Parkinson's disease and in dopamine depleted rodent models (Brown 2007; Jenkinson and Brown 2011). A direct correlation has also been observed between the severity of movement related Parkinson's symptoms and beta power in the STN, and treatment with L-DOPA therapy reduces this beta power while partially alleviating movement deficits (Jenkinson and Brown 2011). Although all of this work has focused on downstream basal ganglia nuclei, other studies in normal rodents have shown that beta bursts tend to be simultaneously expressed throughout the basal ganglia and some connected cortical structures (Leventhal, Gage et al. 2012). Therefore, it is reasonable to postulate that Parkinson's patients with dopamine depletion which extends to the ventral regions of the striatum may exhibit abnormally strong beta oscillations there as well. Our study suggests that a potential consequence of this may be a type of cognitive inflexibility in which patients may experience difficulty in adjusting their behavior in response to changes in positive feedback. Behavioral studies in Parkinson's patients suggest that this may in fact be the case (Frank, Seeberger et al. 2004).

2.3.6 Possible mechanisms of beta burst generation.

This study was unable to determine the mechanism by which striatal beta bursting is generated. The inverse phase preferences of the MSNs and FSIs suggest that perhaps the bursting measuring

in the LFP is generated by local interactions between these cell types, but it is still unclear how they come to transiently oscillate in the first place. One possibility is that the oscillations are generated in an upstream structure to the striatum then passed along directly through synaptic connections. In favor of this explanation is the fact that beta oscillations are highly synchronized across cortico-basal ganglia loops (Leventhal, Gage et al. 2012). However, this explanation is not fully satisfactory, as it simply transfers the question of mechanism to another brain structure. Another possibility is that the transient beta bursts are the product of a highly convergent and synchronous re-setting mechanism. In this case, a strong synchronous input to a structure (or several simultaneously) would transiently allow the cells within that structure to briefly fire together synchronously at their resonant firing frequency (Fig 2.14). In the case of striatal FSIs, that frequency happens to be within the beta range (~15-30Hz). The network would briefly oscillate after the reset event then would quickly shift back into a desynchronized state. There are several candidates that could produce such a synchronous reset, including the cortex and the ventral pallidum, which sends a selective convergent inhibitory back-projection to striatal FSIs (Humphries and Prescott 2010).

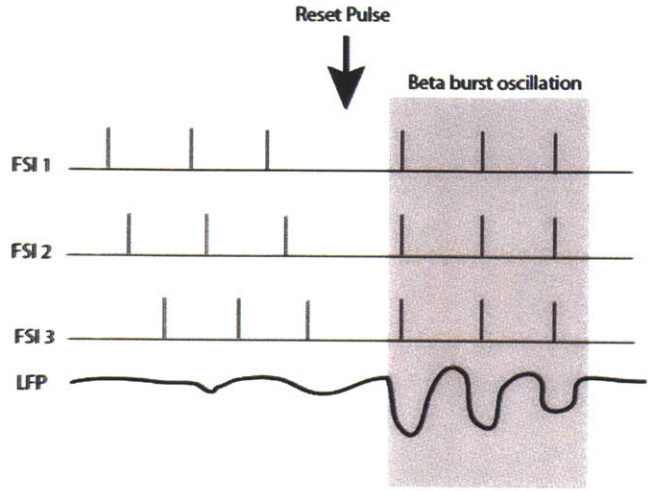

Figure 2.14 Reset mechanism for beta burst generation. Spike trains of FSIs are initially asynchronous but become transiently synchronized following a common pulse of input (in this case, inhibition).

2.4 Methods

Behavioral Task. All training and recording sessions were conducted with the rats on an elevated T-maze described previously (Barnes, Kubota et al. 2005). Photobeam units were placed along the outer walls of the maze, allowing the detection of behavioral events, which were stored with timestamps synchronized with those for the spike and LFP data. Each trial on the T-maze task began with a warning click signaling the beginning of the trial. After the click, a gate was manually swung into place, allowing the animal to run down the central arm of the maze. Approximately half way down the maze, rats broke a photobeam that in turn triggered either a 1 kHz or 8 kHz tone. The frequency of the tone was associated with reward delivered at the end of either the left or right maze arm, and tone-reward contingencies were counterbalanced across rats. Four rats were trained with a chocolate sprinkles reward, and 3 of the rats were trained with ~0.3 cc of chocolate milk (Chug brand) as the primary reward. Preliminary analysis of data for these two reward groups indicated no significant differences in learning rates or behavior, so data from all rats were combined for all analyses reported here. Chocolate milk was delivered through a manual pump after animals completed a turn to the correct arm. For reward-delay trials, milk was delivered through automated pumps that were triggered 2 s after the goal-reaching photobeam was broken. After 10 consecutive days of overtraining, these rats were switched to a tactile version of the task (not reported here).

LFP Data Collection and Analysis. Neural signals sampled at 1 kHz were amplified (gain: 1000) and filtered (1-475 Hz) by Cheetah data acquisition system (Neuralynx, Bozeman, MT). LFPs were referenced to the external ground of the recording system. The spectral content of the

LFP signals was analyzed using open-source Chronux algorithms (http://chronux.org), in-house software, and the Matlab Signal Processing Toolkit (MathWorks, Natick, MA). Spectrograms were constructed by the multi-taper method (DeCoteau, Thorn et al. 2007) with 3 tapers, time bandwidth product of 2, and window width of 0.75s. For computing average power around events, power in the beta band (15-28 Hz) was averaged for each session (across 40 trials) and then was normalized by converting it to z-scores relative to beta power in all other peri-event windows. All sessions for all rats were then combined separately for correct and incorrect trials. For calculating correlations of beta and gamma (70-90 Hz) power with performance, power was first averaged across tetrodes for each session in one of two windows: 3s before goal reaching or 0-1 s after goal-reaching. Z-scores were computed for mean values for each session relative to the set of mean power values in the pre-goal or post-goal window for each individual rat. Normalized values for all sessions were then combined across rats to compute linear correlations with performance. High amplitude beta and gamma oscillations used for spike alignment and synchrony analysis were identified by first band-pass filtering single trial LFP traces in the beta or gamma range, and then their voltage values were converted to z-scores for each trial separately. Peaks were identified where the normalized trace crossed a threshold of >2.5 SD above the trial mean. Peaks were considered members of individual bursts if consecutive threshold crossings were separated by at least one oscillation cycle. Only the largest positive going peak of each identified burst was used for analysis. Phase synchrony across LFPs during beta bursts was assessed by first identifying the tetrode with the highest amplitude peak during each identified burst. This tetrode served as the reference for computing phase differences between LFPs on other tetrodes. For convention, this reference phase was set to 90 degrees. Phases on other tetrodes, relative to the peak time for the peak reference electrode, were

calculated using a Hilbert transform on each band-pass filtered trace. These phase values were used to construct the phase difference histograms in.

Spike Data Collection and Analysis. Signals for spike collection were sampled at 32 kHz then were amplified (gain: 2,000-10,000) and band-pass filtered (600-6,000 Hz) before offline storage. Manual spike sorting for single unit separation was performed using Offline Sorter (Plexon v. 2.8.7, Plexon, Dallas, TX). After sorting, units were classified as putative medium spiny projection neurons (MSNs), fast-spiking interneurons (FSIs), or tonically active neurons (TANs) by manual examination of interspike intervals (ISIs), waveform shape, and firing rates (Kubota, Liu et al. 2009). In general, putative MSNs were easily distinguishable from putative FSI's by their significantly lower baseline firing rates and greater proportion of ISIs >100 ms. TANs had firing rates intermediate to MSNs and FSIs. These made up a small proportion of our dataset and were not included in the analysis.

Significant modulation of single units was determined using circular statistics on phase histograms. To construct phase histograms, instantaneous phases of the band-pass filtered trace for beta and gamma oscillations were determined using a Hilbert transform and spiking was binned by phase (20 phase bins) in a window ±150 ms around the peaks of high amplitude beta and gamma events recorded on the same tetrode as the spikes. The phase distribution for each unit was then tested for uniformity with a Rayleigh test ($P < 0.05$) to identify significant modulation of individual unit spiking by high amplitude beta or gamma bursts. For determining whether spikes align also to phases of LFPs recorded on spatially separated tetrodes, spike-phase histograms were constructed around band-pass filtered LFPs on a randomly chosen tetrode that

was different from the one on which the spikes were recorded. Average phase histograms were constructed by computing z-scores for the binned spike counts of each unit relative to the mean and standard deviation of the bin counts for that unit. The mean and standard error of the z-scores were computed for each bin across all significantly modulated units. Statistical differences between the strength of average z-score normalized phase distributions were computed by performing a Kruskal-Wallis test on the absolute values of the average z-score normalized distributions (local vs. non-local for beta and gamma separately). Spike histograms around burst peaks were constructed by aligning spiking on the peaks of accepted beta and gamma bursts (bin size = 3 ms for beta, 2 ms for gamma). For the population average, firing rates were converted to z-scores relative to the mean and standard deviation of each unit's distribution. Histograms were then averaged across units and scaled to the absolute value of the minimum of the distribution. To test whether spikes from a particular unit were aligned with peaks of bursts on tetrodes that did not record the unit spiking, spikes were aligned to burst peaks from randomly chosen tetrodes. Significant differences between local and non-local histograms were computed by first determining the highest normalized firing rates for each unit around the peak of the beta or gamma burst. The baseline z-score (200-300 ms before burst peak) for each unit was then subtracted from the max z-score for that unit. A two tailed t-test was then conducted to compare the mean of all baseline subtracted max z-scored firing rates in the local condition with the mean of the max z-scored firing rates in the non-local condition for beta and gamma independently.

Histology. Brains were fixed by transcardial perfusion with paraformaldehyde and post-fixed in 4% paraformaldehyde in 0.1M NaKPO4 buffer, and 30 um-thick transverse sections were cut on

a freezing microtome and were stained with cresylecht violet to allow reconstruction of the recording sites.

ACKNOWLEDGMENTS. This work was funded by NIH Grant R01 MH060379 and a Mark Gorenberg graduate student fellowship. We would like to thank Christine Keller-McGandy, who was responsible for performing the histology, and Henry F. Hall.

Chapter 3: Dopamine dynamics in the basal ganglia during motivated pursuit of goals

3.1 Introduction and background

The dopamine system in the midbrain broadcasts modulatory signals to large regions of the mammalian brain (Fig 1.4). These signals are critical for controlling motivational drive and effort to reach goals and for learning about external stimuli that can reliably predict those goals (Schultz, Dayan et al. 1997; Salamone and Correa 2002; Schultz 2002; Berridge 2007; Salamone, Correa et al. 2007). Neurological conditions that affect the dopamine system result in complex deficits in both motor and cognitive domains. In Parkinson's disease, which is marked by a slow degeneration of dopamine neurons in the Substantia Nigra pars Compacta (SNc) patients experience severe problems in initiating movements and in maintaining drive to complete those movements. They also show deficits in some forms of learning (Frank, Seeberger et al. 2004). Addiction, unlike Parkinson's disease, is believed to result from an abnormal overactivity of dopaminergic systems, which results in pathologically compulsive behaviors (Redish, Jensen et al. 2008; Schultz 2011).

Electrophysiological studies of dopamine neuron activity in behaving animals have provided some insight into the mechanisms underlying the learning and motivational functions of the dopamine system. Neurons in the dopaminergic ventral tegmental area fire action potentials in two distinct modes, termed "tonic" and "phasic." Phasic firing occurs in bursts of ~60Hz and is

dependent on post-synaptic NMDA receptors (Grace and Bunney 1984; Overton and Clark 1992; Zweifel, Parker et al. 2009), while ongoing tonic firing occurs at ~5Hz and is dependent on the activity of post-synaptic voltage gated Ca+ channels (Grace and Bunney 1984; Shepard and Bunney 1991). A large body of work has focused on the properties of phasic VTA firing in the awake, behaving animals. Phasic bursts of activity have been found to occur in response to unpredicted rewarding events, and across learning, as the rewards become more predictable, phasic firing shifts to the earliest cues that predict that reward (Schultz, Dayan et al. 1997; Schultz 2002). When an expected reward is withheld, the VTA neurons show a dip in their firing rates. These findings have led to the highly influential idea that dopamine neurons encode information about "reward prediction errors" that can be used to modify strengths of synaptic connections and guide reinforcement learning (Schultz, Dayan et al. 1997; Reynolds and Wickens 2002; Schultz 2002; Wickens, Horvitz et al. 2007). Another complementary interpretation of these findings is that phasic dopamine signals in response to reward predictive cues provide an immediate motivational drive to initiate reward seeking behaviors (Berridge 2007).

Classic phasic dopamine signaling alone cannot account for crucial role of dopamine signaling in controlling ongoing effort and motivation, however. If dopamine signaling through D2 receptors is blocked in rats asked to choose in a T-maze between a pursuing a large reward, but at a larger effort cost, or a smaller reward with a smaller effort cost, rats will bias their choices towards the lower effort option (Salamone, Correa et al. 2007). Moreover, transgenic mice with phasic signaling selectively attenuated in dopamine neurons exhibit deficits in some forms of learning and cue-guided movements, but show no impairments in general motivation (Darvas and

Palmiter 2009; Zweifel, Parker et al. 2009). It is likely, therefore, that slower timescale fluctuations in dopamine signaling play an essential role in determining levels of ongoing behavioral motivation.

Until recently, the only method of studying the dynamics of dopamine signaling on a behaviorally relevant timescale was to record neural activity from dopaminergic nuclei such as the VTA and SNc. The problem with this technique is that ~50% of cells in these nuclei are GABAergic and do not release dopamine (Margolis, Lock et al. 2006; Nair-Roberts, Chatelain-Badie et al. 2008; Cohen, Haesler et al. 2012). To circumvent this issue, cell specific targeting of dopamine and GABA cells with light sensitive channelrhodopsin has been employed to allow cell-type verification prior to recordings in behavioral experiments (Cohen, Haesler et al. 2012). An alternative to these methods is fast-scan cyclic voltammetry (FSCV), which allows dopamine release in target sites to be measured directly on subsecond timescales. One advantage of this technique is that it can detect regional heterogeneities in dopamine signaling. Moreover, dopamine terminals are under the influence of considerable pre-synaptic modulation, particularly in the striatum (Krebs, Trovero et al. 1991; Zhang and Sulzer 2003; Calabresi and Di Filippo 2008; Exley, McIntosh et al. 2012), so electrophysiological recordings of dopamine neurons may not fully capture the true dynamics of dopamine signaling in target structures. We used FSCV to study changes in dopamine signaling in the striatum during acquisition and overtraining of the T-maze task described in Chapter 1. Our goals were several-fold and are outlined below.

First, we wanted to determine the dopamine signaling patterns for behaviors that require extended navigation through space. Several studies have implicated dopamine signaling in the

striatum in learning and performing spatial navigation behaviors such as the Morris water maze (Whishaw and Dunnett 1985; Braun, Graham et al. 2012), but the dopamine dynamics during these tasks is unknown. All prior studies of dopamine neuron firing and dopamine release have focused on instrumental or classical conditioning tasks which require animals to perform a specific action to retrieve reward (such as lever pressing) or to associate a particular cue (such as a tone) with reward delivery (Fig 3.1). These behaviors are carried out in a fixed behavioral context, such as an operant chamber for rats, and the requirements to perform the tasks are fundamentally different from those that require an animal to track a distant goal through a changing environment. Most notably, spatial navigation requires that ongoing motivational levels and behavioral state be constantly updated depending on the animal's location in the environment. If the animal finds itself in a place they don't recognize, or (having made a wrong turn) a place that is farther from their goal than predicted, they might need to reassess their behavior and change their direction of movement. Thus, spatial navigation tasks may be ideal for detecting the effects of changing context on dopamine signals that control ongoing motivational levels.

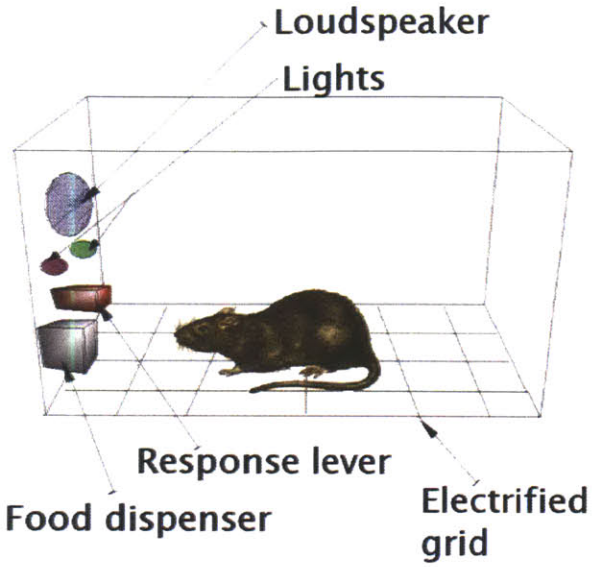

Figure 3.1 Standard chamber for operant and classical conditioning tasks. Operant tasks require the rat to press a lever in response to a visual or auditory stimulus to receive a food reward at the dispenser. Classical conditioning tasks do not require a lever press action to trigger the reward delivery in response to the stimulus. Image from http://en.wikipedia.org/wiki/Operant_conditioning_chamber

Second, we wanted to determine how dopamine signaling in the T-maze task changes with learning and with overtraining. Our lab and others have identified learning related plasticity in neural firing and in local field potential oscillations in the striatum during similar T-maze tasks (DeCoteau, Thorn et al. 2007; van der Meer and Redish 2009; Howe, Atallah et al. 2011). Dopamine signaling has been implicated as a major contributor to striatal synaptic plasticity, so we reasoned that changes in the patterns of dopamine signaling during learning may correlate with the striatal neural activity. To explore direct relationships between striatal dopamine release and neural activity during T-maze learning, we developed a new method for performing FSCV and tetrode recordings simultaneously from nearby sites (Fig 3.3).

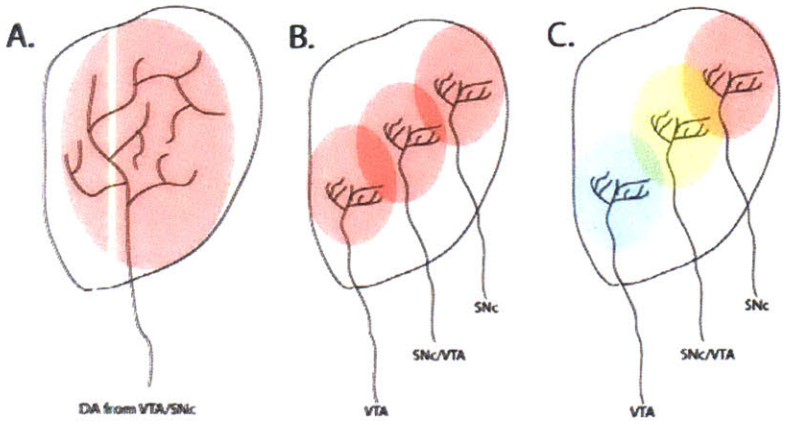

Figure 3.2 Illustration of different possibilities for the mappings of dopamine projections and release patterns. **A.** Broad projections and homogeneous release into the striatum (no functional or structural heterogeneity). **B.** Segregated projections but no functional heterogeneity. **C.** Segregated projections and functional heterogeneity of firing patterns produce heterogeneous patterns of striatal dopamine release.

Finally, FSCV allowed us to test whether dopamine signaling during T-maze behavior varies across different striatal subregions. The issue of whether dopamine is homogeneously broadcast across its target sites or whether there are functional heterogeneities in dopamine signaling has been highly controversial (Fig 3.3). Tracing studies of single dopamine neuron axons have revealed large arborizations that can extend across wide regions of the striatum (Fig 3.2 A; (Matsuda, Furuta et al. 2009). Despite this divergence, a projection topography exists: dopamine neurons in the substantia nigra pars compacta (SNc) project to dorsal and lateral regions of the striatum, while dopamine neurons in the neighboring ventral tegmental area (VTA) project mainly to ventral and medial striatum (Haber, Fudge et al. 2000; Joel and Weiner 2000). Electrophysiological and electrochemical studies have reached conflicting conclusions about

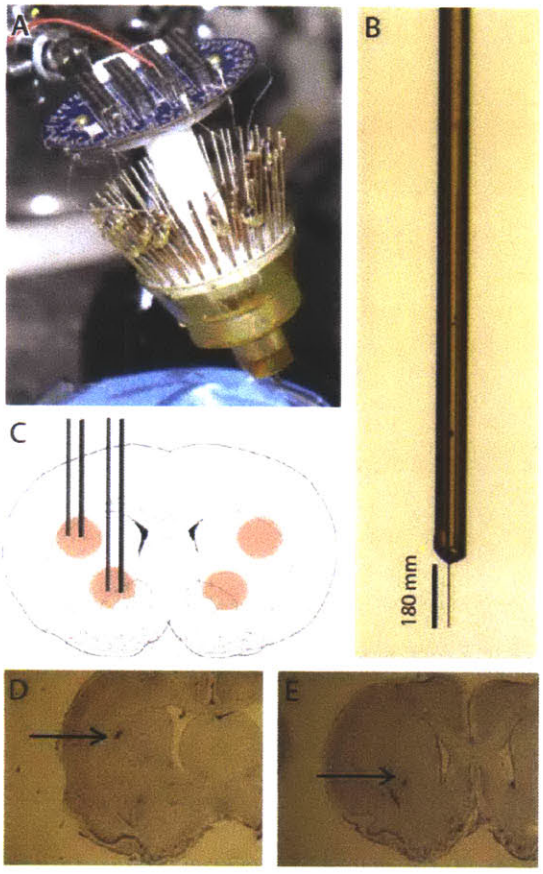

Figure 3.3 Method for multi-site chronic fast-scan cyclic voltammetry (FSCV) recordings from the DLS and VMS. A. Implantable headstage recording device loaded with independently movable carbon fiber probes for striatal dopamine recordings. **B.** Chronically implantable FSCV electrode from our lab consists of a carbon fiber electrode enclosed within a flexible glass silica tube. **C.** Coronal schematic of a rat brain showing FSCV recording sites in the DLS and VMS. **D-E.** Nissl stained coronal rat brain section showing lesions marking location of voltammetry probes in the DLS (**D**) and VMS (**E**).

whether dopamine neuron signaling is uniform across its target regions. Some recording studies in dopaminergic nuclei have revealed a high degree of homogeneity in dopamine neuron responding (Morris, Arkadir et al. 2004), while other studies have shown that neurons in neighboring dopamine nuclei can respond to stimuli in different ways: some responding to salient events in general regardless of valence and others responding only to positive rewarding events suggesting the possibility of functional heterogeneity of dopamine release (Matsumoto and Hikosaka 2009). Further support for functional heterogeneity in the striatum has come from FSCV in behaving animals during Pavlovian tasks (Brown, McCutcheon et al. 2011). This question has important implications for understanding how downstream striatal regions are influenced by dopamine signaling.

3.2 Results

3.2.1 Validation of dopamine signals measured with FSCV.

We outfitted 9 rats with headstages (Barnes, Kubota et al. 2005), each carrying independently movable microsensor probes for FSCV recordings, two in the dorsolateral striatum (DLS) and two in the ventromedial striatum (VMS), and we measured dopamine release to investigate real-time striatal dopamine signaling as the rats navigated mazes of different size and shape to retrieve reward. Dopamine release was initially verified by electrical stimulation of the medial forebrain bundle under anesthesia (Fig 3.4 A, B, and C). Using the templates of stimulated dopamine release, we were able to convert measurements taken during maze running into relative measures of dopamine concentration using a chemometric technique called principle

component regression analysis (Keithley, Heien et al. 2009). Prior to and following selected behavioral sessions on the maze, we measured the dopamine transients elicited by unexpected rewards and salient light flashes to verify that the recording sites in the striatum supported dopamine release (Fig 3.4 D, E, and F). In a separate subset of rats (n = 3), we verified that the MFB stimulation paradigm we were using was specifically facilitating the release and detection of dopamine release with the FSCV (Fig 3.5). These rats received unilateral injections of 6-OHDA into the DLS, and striatal FSCV and MFB stimulation electrodes were implanted bilaterally. Stimulation in the non-lesioned hemisphere produced robust dopamine release, while stimulation in the lesioned hemisphere produced very little detectable release, verifying that the MFB stimulation reliably produced local dopamine release, detectable by FSCV.

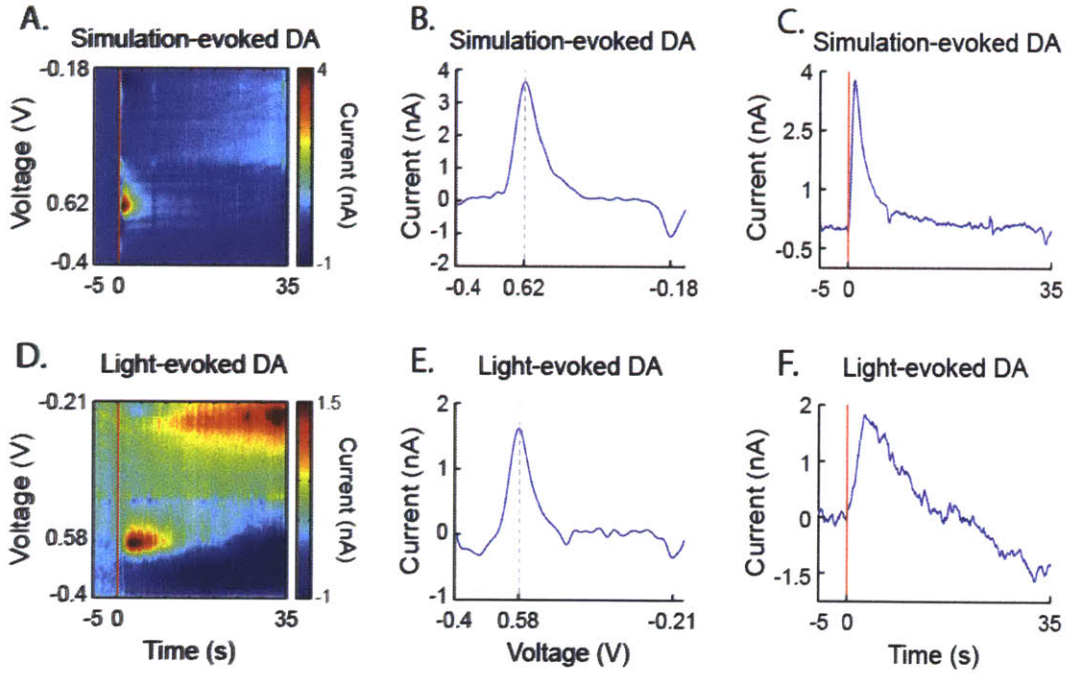

Figure 3.4 Light-evoked dopamine release in the VMS of behaving rats verified by electrical stimulation of the medial forebrain bundle (MFB). A. Color plot showing measured current at each potential during successive scans with electrical MFB stimulation (red line) in an anesthetized rat. **B.** Current-voltage plot from the 5 scans following (0.5 s) MFB stimulation. The large increase in current around 0.6 V (dotted line, peak) corresponds to the dopamine redox potential measured *in vitro*. **C.** Time-course of current at the dopamine redox potential around electrical stimulation (red dotted line). **D-F** Plots, as in **A-C**, illustrating dopamine response in the same rat, now awake, to room lights being turned on.

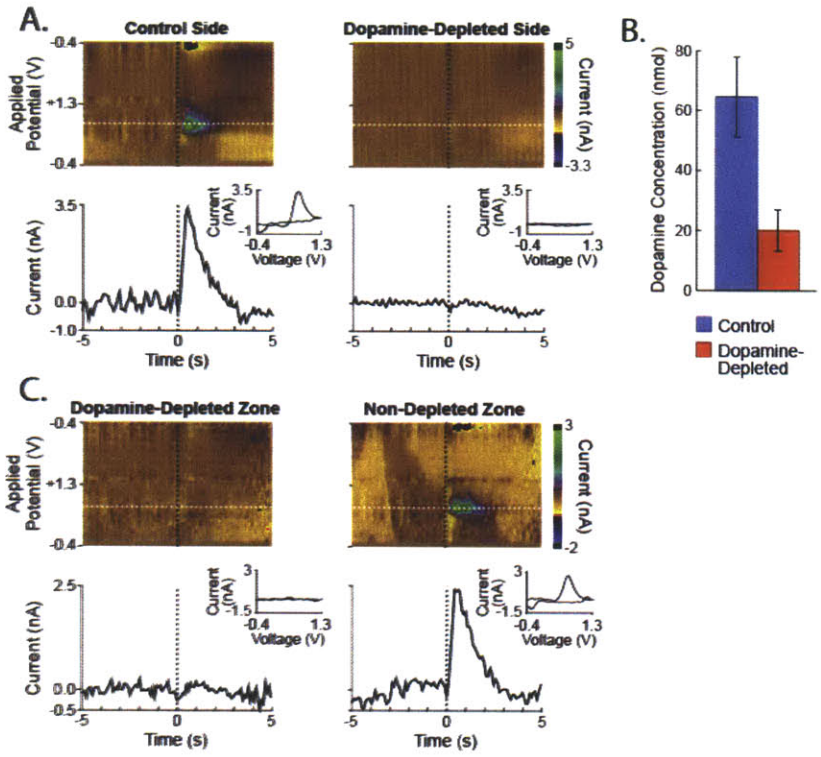

Figure 3.5 Dopamine depletion measured with fast-scan cyclic voltammetry. **A.** Top: Redox currents as a function of potential during consecutive voltammetric scans around the time of MFB electrical stimulation (time 0) on the control side (left) and lesion side (right). Note the increase in current around the peak dopamine oxidation potential (0.6V, white dashed line). Bottom: Currents generated by dopamine oxidation (bottom) showing release on the control side (left) and its absence on the 6-OHDA lesion side (right). Insets show background subtracted cyclic voltammograms after MFB stimulation. **B.** Concentration of dopamine following MFB stimulation in the control and dopamine-depleted striatum. **C.** Phasic dopamine release at a site in the dopamine-depleted region in the dorsolateral striatum (left) and a site ventral to the depleted zone in the same hemisphere (right) in response to MFB stimulation, shown as in **A**.

3.2.2 Dopamine signals ramp from the start of maze running to goal reaching.

The rats were first trained on a T-maze task that required them to start runs after a click cue and then to turn left or right as instructed by an auditory tone to receive a chocolate

milk reward delivered at the correct end-arm (Barnes, Kubota et al. 2005) (n = 9, Fig. 3.6). To our surprise, instead of mainly finding isolated dopamine transients at the initial cue or at goal-reaching, we primarily found gradual increases in dopamine signal that began at the early cue and ended after goal-reaching (Fig. 3.6). These ramping dopamine responses were present in both DLS and VMS recording sites, were more frequent and larger ventrally (Fig. 3.7), and were evident both in single trials (Fig. 3.6) and in population averages (Fig. 3.8). Before goal-reaching, the ramps

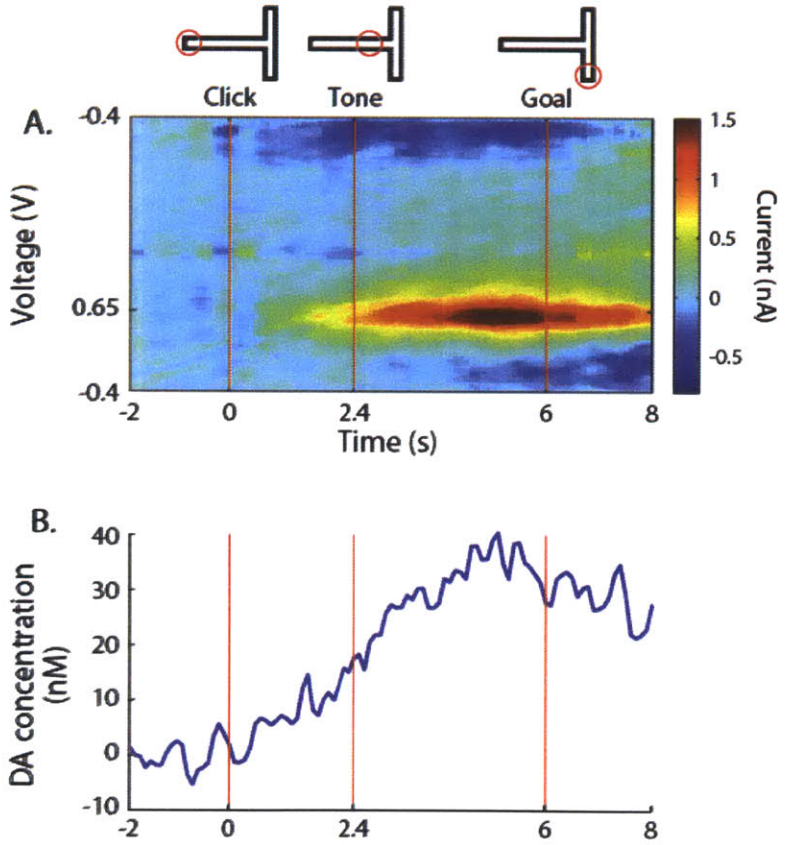

Figure 3.6 Ramping striatal dopamine during maze runs. A. Current as a function of voltage and time relative to warning click onset in the T-maze. Note the gradual increase in current at the dopamine redox potential (~0.6V). Red lines indicate the warning click, tone cue and goal reaching time points. B. Dopamine concentration measured by FSCV in VMS during the trial plotted in A.

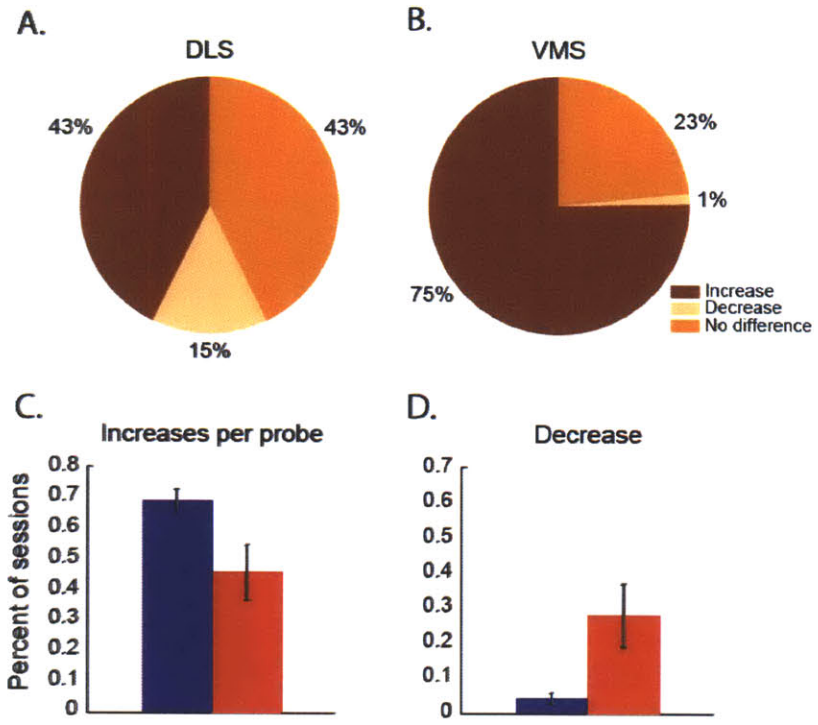

Figure 3.7 Distribution of response types in the DLS and VMS. A. Proportions of recordings in the DLS with a significant (*t*-test, $P < 0.05$) increase in dopamine concentration during a 0.5s window prior to goal-reaching (134/314), decrease (46/314), or no significant change (134/314). **B.** Response types, as in **A,** for VMS recordings: increase (291/387), decrease (5/387), and no change (91/387). **C and D.** Average percent of sessions that showed a significant increase (C) or decrease (D) per microsensor placed in the VMS (blue) and DLS (red). Errorbars, S.E.M.

were similar in amplitude for correct and incorrect trials, though were slightly larger for correct trials; after goal-reaching, the signals were significantly larger in correct trials (Fig. 3.8A). A subset of the recordings in the DLS (15%, 46/314, from 7 probes across 5 rats) exhibited inhibitory ramping to goal-reaching (Fig 3.7 and 3.8C). Such negative ramps were rare in the VMS (< 2%, 5/387 recordings), suggesting that DLS dopamine signals exhibited a larger degree of heterogeneity in polarity than did VMS signals (Fig 3.7A).

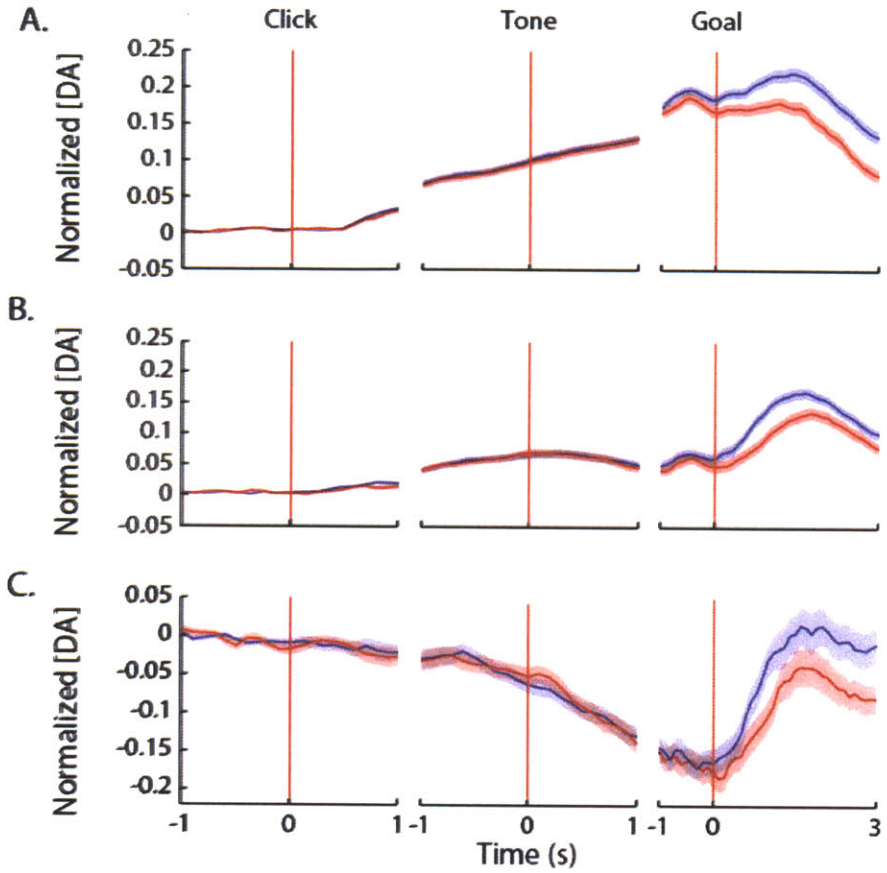

Figure 3.8 Average normalized dopamine traces for DLS and VMS recordings. Normalized dopamine concentration averaged for all VMS **(A)** and DLS **(B)** recordings (n = 387 and 314, respectively), for correct (blue) and incorrect (red) trials. Shading, SEM. **C.** Average normalized dopamine concentration for DLS recordings (n = 46 in 5 rats) with significant (t-test, $P < 0.05$) decrease in dopamine concentration during maze runs.

3.2.3 Phasic transients were present on some maze sessions and were distinct from the ramping responses.

We occasionally identified isolated phasic transients at warning click or at goal-reaching in single trials in the VMS and DLS (Fig 3.9A). The average proportion of trials per recording session containing such transients was low (~8% around click and ~4% after goal-reaching). Isolated phasic transients may have been difficult to identify because most of these phasic signals

were likely superimposed on the slower ramping signals (Fig 3.10). At peak, the magnitudes of the ramping dopamine signals were comparable to those of the isolated phasic dopamine signals recorded here (Fig. 3.9 and 3.10) and in other studies (Day, Roitman et al. 2007; Gan, Walton et al. 2010). These results suggested that the ramping striatal dopamine signals represent a novel form of dopamine signaling during spatial navigation that is distinct from previously described dopamine transients elicited in response to cues and rewards.

3.2.4 Ramping dopamine signals displayed consistent biases towards one of the two maze arms.

Remarkably, when we compared the dopamine signals from runs to the left and right end-arms of the T-maze, we found that that dopamine ramps often were larger for one of the end-arms (Fig.

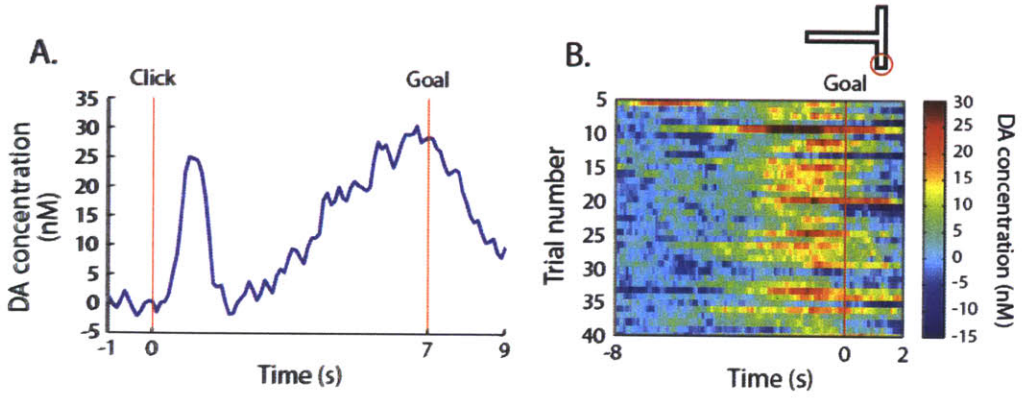

Figure 3.9 Phasic and ramping dopamine signals occur in single trials A. Dopamine concentration in trial with both phasic dopamine response to warning click and slower ramping signal. **B.** Trial-by-trial changes in dopamine concentration over 40 trials, plotted relative to goal-reaching.

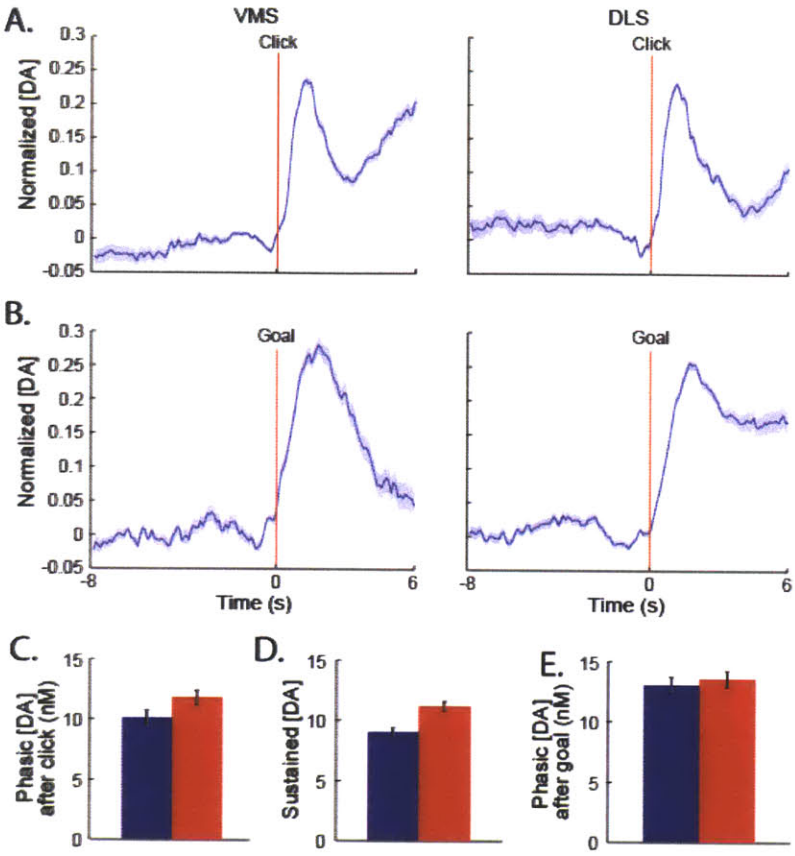

Figure 3.10 Transient responses to cues and rewards are present in the T-maze and are distinct from the sustained ramping response. A. Average normalized dopamine traces from VMS (left) and DLS (right) microsensors from all trials that showed a significant increase in a 0.5 s window after warning click (red bars). Note the sharp increase in dopamine after the click which rides on the slower ramping signal. **B.** Average normalized dopamine as in A for all trials that showed a significant increase following goal reaching but no significant increase prior to goal reaching (see Methods). **C.** Average maximum peak dopamine concentrations for the phasic click transients plotted in A. The blue bar are averages from DLS recordings and the red are averages from VMS recordings. **D.** Average maximum dopamine concentrations for the ramping dopamine signals prior to goal reaching plotted as in C. **E.** Average max dopamine concentrations for the postgoal responses plotted in B.

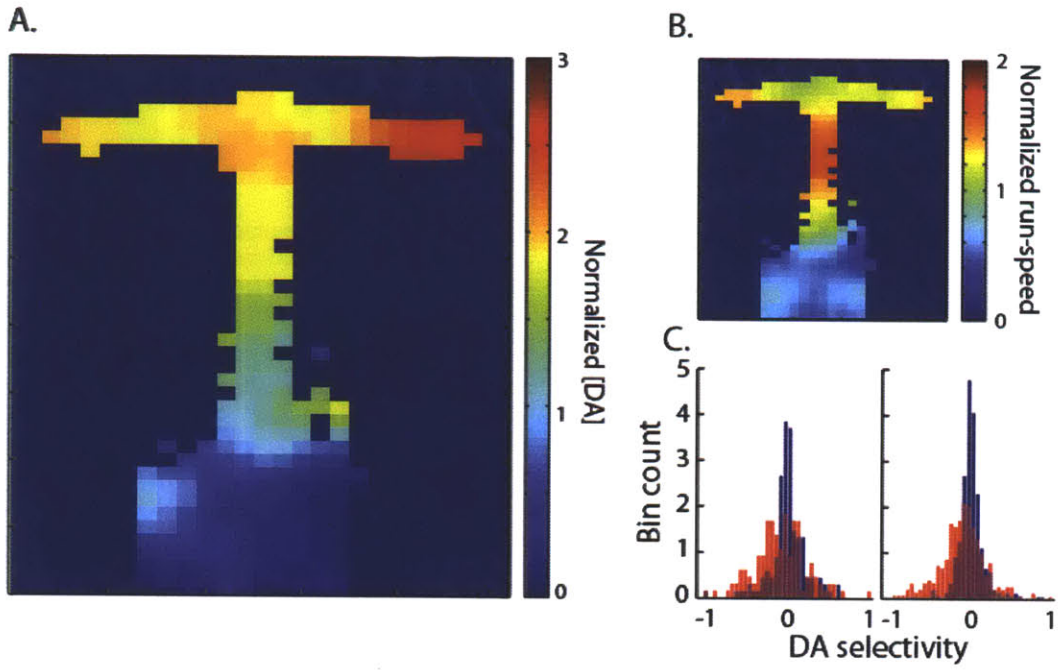

Figure 3.11 Dopamine ramping is selective for maze arm. A and B. Average normalized dopamine at a VMS site (A) and average running speed (B) in one representative VMS recording as a function of maze location (n = 19 sessions). **C.** Distribution of end-arm selectivity indices (see Methods) for all VMS (left) and DLS (right) recordings (red) compared to shuffled data (blue) for all rats (n = 9). Note the larger variance in the non-shuffled data, indicating significant end-arm biases in the dataset.

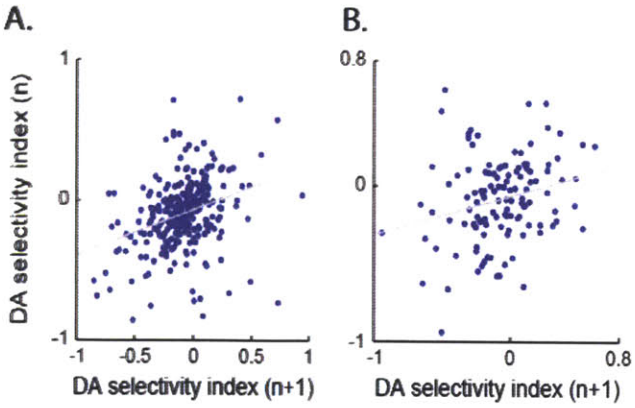

Figure 3.12 Dopamine end-arm preferences are maintained across consecutive sessions. Pairs of selectivity indices in consecutively recorded sessions in the VMS (**A**) and the DLS (**B**) are positively correlated (Pearson's r = 0.32, P = 1.2e-7; r = 0.25, P = 0.0064; respectively)

3.11). Such arm preferences were often stable across multiple 40-trial recording sessions (Fig. 3.11 and 3.12), and they were observed significantly more often than chance in both the DLS and the VMS (Fig. 3.11C, z-test, $P < 0.00001$ vs. bootstrapped variances). These preferences emerged gradually with training as task performance improved (ANOVA, $P = 0.032$; Fig. 3.15), and as training progressed these dopamine selectivities tended to match the behavioral biases that we found to develop in the task, favoring the right end-arm, regardless of whether the dopamine ramps were recorded in the right or left hemisphere (Fig. 3.15, C and D). This increase in end-arm selectivity across training was not accompanied by an increase in magnitude of the ramping signals (data not shown). These observations suggested that end-arm biases of the extended dopamine signals were dependent on repeated exposure to the two maze end-arms and could modulate or reflect learned biases in behavioral choices.

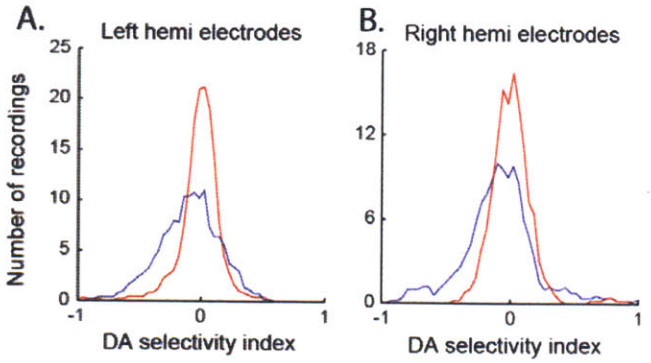

Figure 3.13 Right arm bias in the extended dopamine signal is present in both left and right hemisphere recordings. A. Distribution of selectivity indices as in Fig. 2c for all probes implanted in the left hemisphere ($n = 4$ rats) and in the right hemisphere **(B)** ($n = 3$ rats). Note the bias in both groups of the selectivity preference towards negative selectivity indices (right bias, blue) relative to the shuffled data (red).

These end-arm preferences of the dopamine signals were mostly similar for different recording sites in the DLS and VMS (data not shown), but we did observe significant end-arm biases that were opposite for different probes in the same hemisphere in ten sessions (~5% of sessions recorded in six of the nine rats), and in nine of these ten sessions, one probe was in the DLS and the other was in the VMS. Moreover, in one rat, we found consistent opposite end-arm preferences for DLS and VMS probes of one hemisphere across multiple consecutive sessions

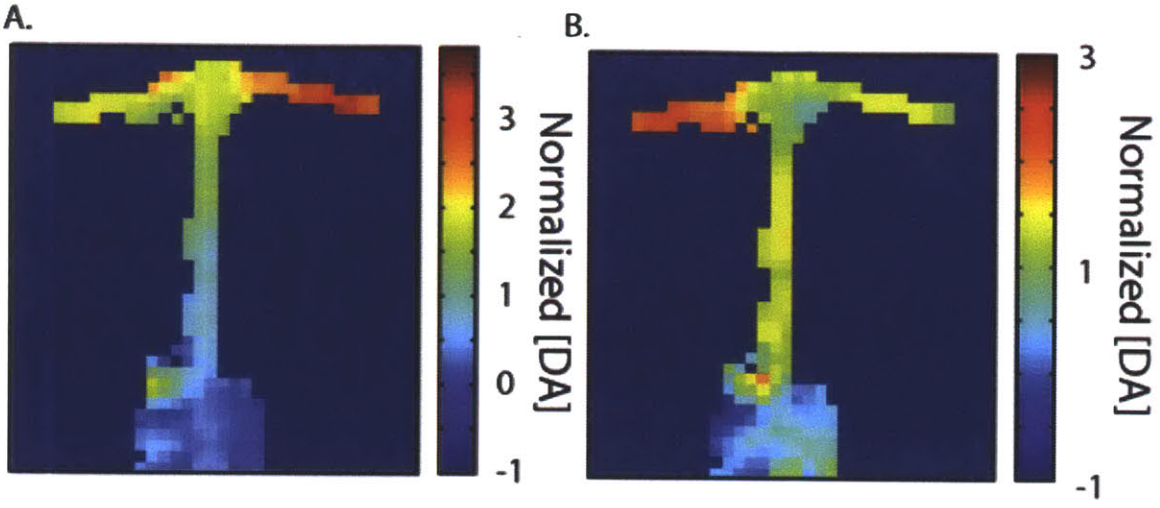

Figure 3.14. Dopamine arm preference heterogeneity is occasionally present across 2 different probes in the same rat. Dopamine averaged across 11 sessions measured simultaneously from probes in DLS (A) and VMS (B).

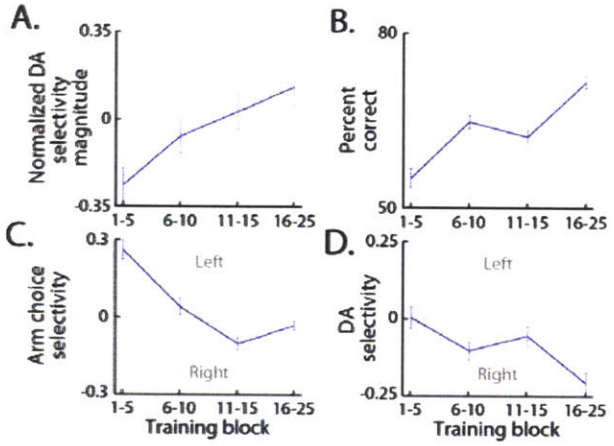

Figure 3.15 Training dependent effects on ramping dopamine selectivity. A. Absolute magnitudes of all z-score normalized selectivity indices (see Methods) averaged within session blocks defined as indicated. Selectivity increases gradually across training. **B.** Percent-correct performance averaged within session blocks as in A. **C.** Arm choice selectivity. **D.** Signed magnitude of selectivity indices averaged for consecutive training blocks. Error bars, SEM. Arm choice selectivity matches dopamine selectivity with extended training.

(Fig. 3.14). These results raise the possibility of heterogeneity in the distribution of the spatially selective signals.

3.2.5 Ramping dopamine signals were sensitive to manipulations of reward value and scaled with changes in maze length.

Given that phasic responses of dopamine-containing neurons can reflect the relative value of stimuli (e.g., (Tobler, Fiorillo et al. 2005), we asked whether the biases in the ramping dopamine signals could be modulated by the size of the rewards delivered at the goal sites. We introduced "M"-shaped (n = 3 rats) and "S"-shaped (n = 2 rats) mazes to test systematically whether the amplitudes of the dopamine signals could be manipulated by reward size, and if so, whether such value-sensitive effects would extend over longer distances (Fig. 3.16). In the M-maze, we gave rats initial 2-4 day exposures to the locations of the asymmetric rewards. We observed robust dopamine ramping that scaled flexibly with the longer arm distance and strongly favored the end-arm with the larger reward (Fig. 3.17 and 3.18). We then reversed the locations of the small and large rewards, and found that the ramping signals also shifted to favor the new high-reward end-arm (Fig. 3.18). We observed similar value-sensitivity of the extended dopamine signals when asymmetric rewards were given on the standard T-maze (Fig 3.17). Run speed profiles and run times in the two end-arms were roughly equivalent (Fig. 3.18C), suggesting that elapsed time to goal-reaching or locomotion differences did not account for these large differences in amplitude of the ramping dopamine signals. The shifts in the dopamine signals following reversal were implemented within the two sessions and were stable thereafter.

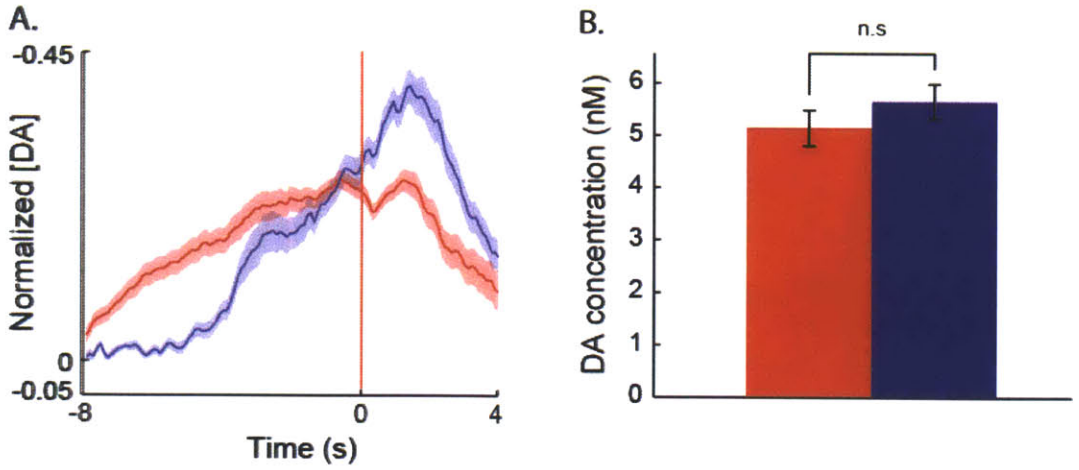

Figure 3.16 Ramping dopamine scales with distance to goal location when maze is lengthened. A. Average normalized dopamine signals relative to goal reaching (red line) from 2 rats on the last 3 T-maze sessions (blue, $n = 54$ recordings) and on the M-maze sessions (red, n = 69 recordings). Dopamine signals peak at similar levels at goal reaching despite the longer distance traveled in the M-maze. **B.** Average peak dopamine concentrations on the T (blue) and M (red) mazes are not significantly different (t-test, $P = 0.56$).

3.2.6 Ramping dopamine signals were independent of the specific actions taken to get reward.

The "S"-shaped maze allowed us to ask whether such value-related differences in the ramping dopamine signals would occur when the actions required to reach the distant goal sites were equivalent (i.e., two left turns followed by two right turns to reach the goal in either direction). The rats had to shuttle back and forth between goal sites to retrieve either a large or small reward at each goal site. Robust ramping dopamine signals were present over the longer distances of the S-maze (Fig. 3.19) with linearity similar to that observed in the shorter mazes. They were larger for the run-trajectories leading to the larger rewards, despite the fact that the sequence of turns

and lengths of the runs needed to reach the larger and smaller rewards were equivalent for both trajectories (Fig. 3.19).

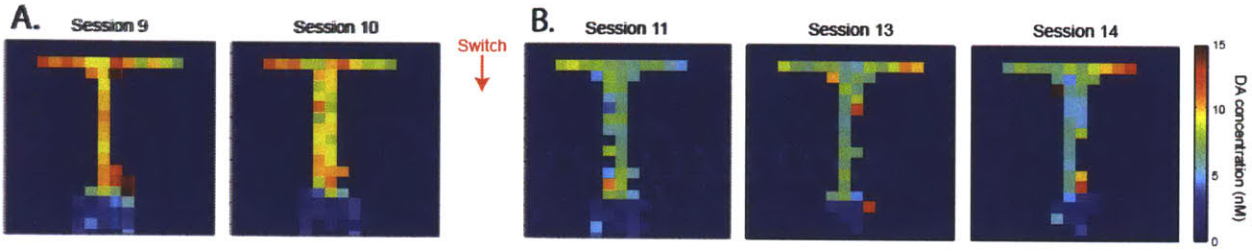

Figure 3.17 Ramping dopamine signals are sensitive to the size of the reward delivered at the maze goal in the forced-turn T-maze. A. Dopamine signals averaged across trials recorded from one probe in the VMS for consecutive 40-trial sessions in which the amount of chocolate milk reward was larger in the left end-arm (0.4 cc) than in the right end-arm (0.1 cc). **B.** Dopamine signals, as in **A**, for three consecutive sessions after reward amounts were reversed between the two end-arms.

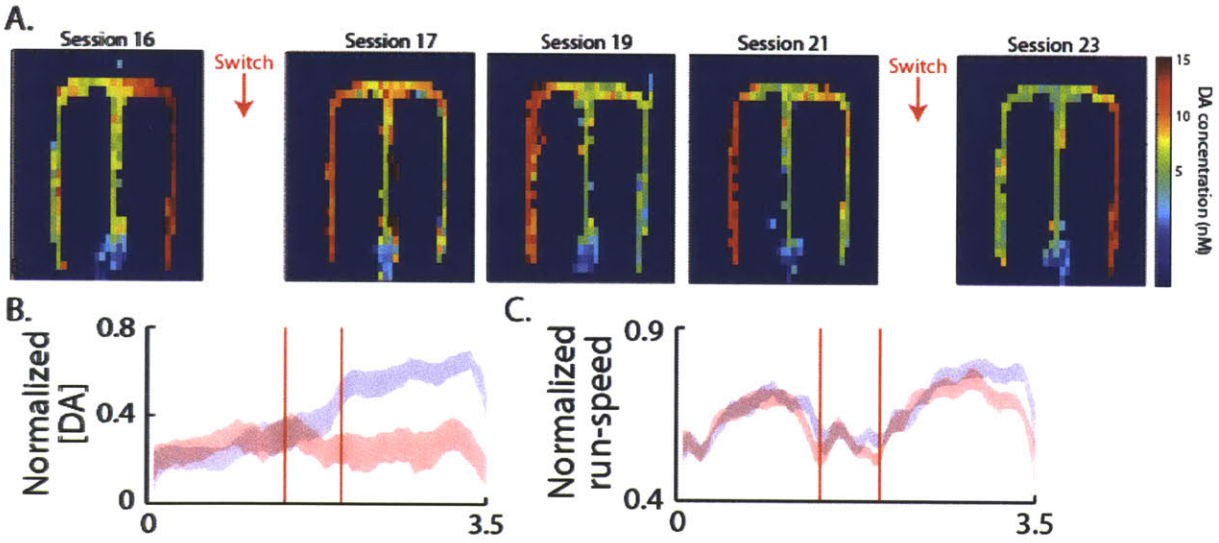

Figure 3.18 Dopamine ramping is sensitive to reward magnitude in an extended M-maze and extends to the termination of maze running. A. Dopamine signals from a VMS probe, for consecutive sessions with larger reward at the right end-arm (session 16) following reversal in reward amounts (17, 19 and 21) and second reversal (23). **B** and **C**. Average normalized dopamine (B) and running speed (C) (n = 17 sessions in 3 rats) as a function of spatial location (x axis, meters). Blue and red, runs toward high and low reward arms. Vertical markers indicate turns. Shading, SEM.

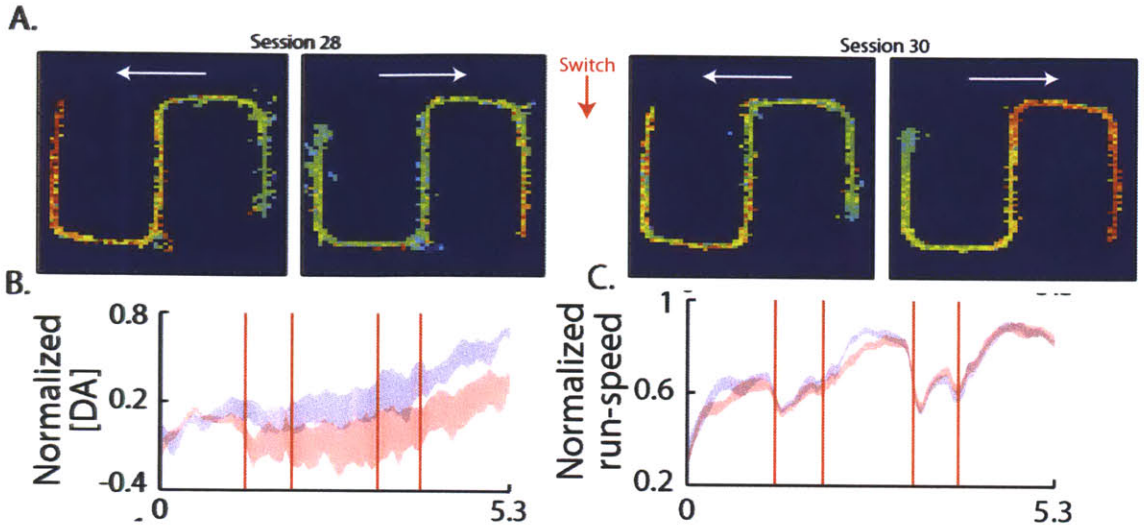

Figure 3.19 Dopamine is signals are selective for trajectory associated with larger rewards on an extended S-maze. Rats were required to run back and forth along the two trajectories to receive a high (0.4cc) or low (0.1cc) quantity of chocolate milk. **A.** Average VMS dopamine in sessions with higher reward at left (session 28) and right (30) goals. Arrows indicate run direction. **B and C.** Average normalized dopamine (B) and running speed (C) for all recordings (n = 9 sessions in 2 rats). Shaded regions, S.E.M. X axis, meters. Red lines indicate turn locations in the S-maze.

These findings demonstrated (1) that extended, ramping dopamine signals can be maintained over distances from reward ranging from 1.5 m (T-maze) to 5.3 m (S-maze), often without detectable accompanying phasic responses, (2) that these prolonged dopamine signals occur whether instructed decisions are required (T-maze) or not (M- and S-mazes), and (3) that the extended dopamine signals scale with distance and reward value even when there are equivalent egocentric requirements for goal-reaching (S-maze).

3.2.7 End-arm biases in dopamine signaling were inversely correlated with beta power at task end.

To ask whether these biases in the ramping dopamine signals would be reflected in simultaneously recorded neural activity in the striatum, we made concurrent chronic measurements of dopamine signals and neural activity from nearby VMS regions in 4 rats trained on the T-maze task. Based on our previous VMS tetrode recordings showing that beta-band oscillations dominate local field potential activity at goal-reaching on correctly performed trials in the T-maze task (Howe, Atallah et al. 2011), we focused our analysis on this frequency band, and we again observed strong post-goal beta-band activity on correct trials (Fig. 3.20A). We tested for, and found, end-arm biases in the power of the post-goal beta signals (Fig. 3.20A and E). Strikingly, these were negatively correlated with the arm preferences of the dopamine ramps that were present before and after goal-reaching (Fig. 3.20E; $r = -0.51$, $P = 0.002$). This inverse relationship could not be accounted for by end-arm differences in reward size, as the negative correlation was observed on correct trials when animals were receiving the same reward amounts in the two end-arms. These findings thus suggested that ramping dopamine signaling is integrated with the on-going neuronal activity dynamics of the VMS (Howe, Atallah et al. 2011).

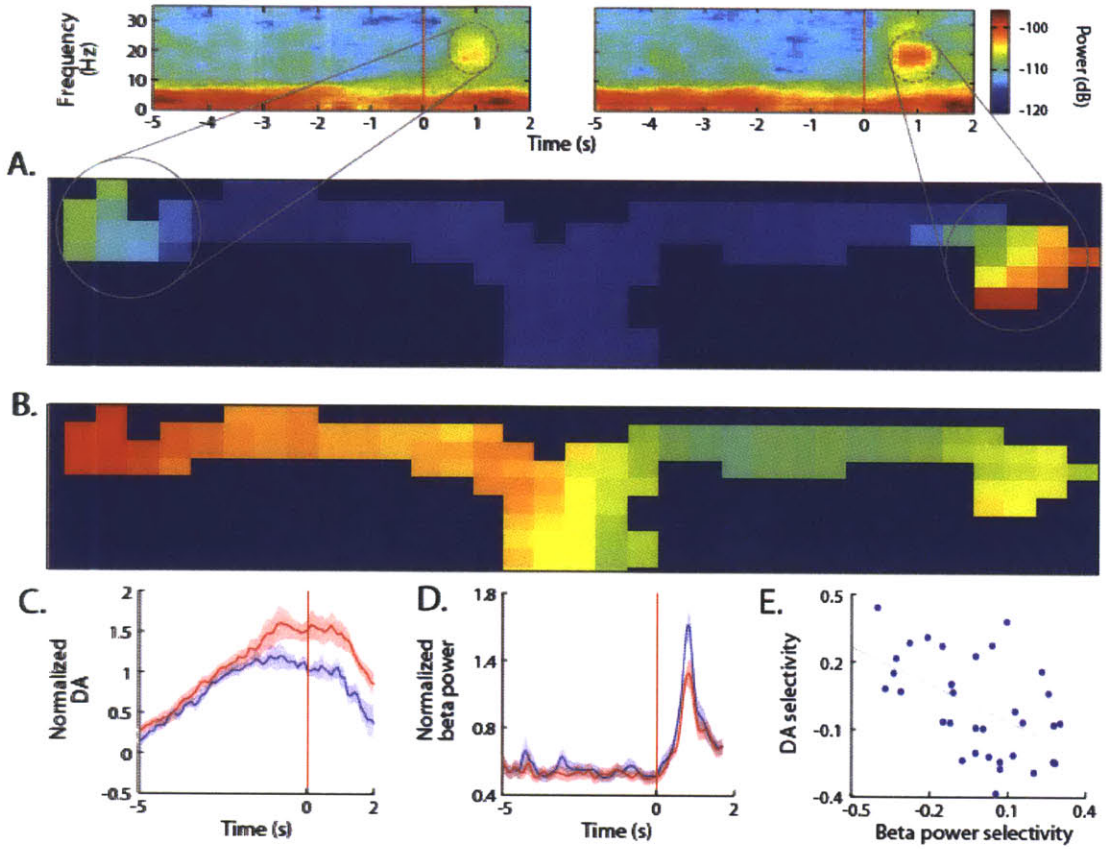

Figure 3.20 Ramping dopamine and beta LFP oscillations have inverse spatial selectivities. A and B. Average beta power (A) and dopamine (B) recorded simultaneously in the VMS in a T-maze session. Insets in (A) show LFP power spectrograms for correct trials to each end-arm. **C.** Average normalized dopamine in VMS during all sessions with simultaneous tetrode recordings (n = 4 rats, 34 sessions) to end-arms with weaker (red) and stronger (blue) beta power. Shading, SEM. **D.** Average normalized beta power for trials to end-arm with higher (red) and lower dopamine (blue). **E.** Ramping dopamine selectivity is negatively correlated with beta power selectivity on correct trials for all recorded sessions (Pearson's r = −0.51, P = 0.002).

3.3 Discussion

In this study, we have described the dynamics of dopamine signaling and oscillatory neural activity in the striatum of rats learning to pursue distant rewards in maze tasks. We revealed a novel type of dopamine signal that is unlike the classic phasic transients to cues and rewards measured here and elsewhere. A primary characteristic of this signal was that it ramped steadily as the rats approach the goal location in the maze. The length of the ramping scaled in proportion to the size of the maze environment, so that the slope of the ramping was steeper in smaller mazes. Interestingly, the ramping dopamine signals often displayed strong selectivity for one of two routes through the 2-arm mazes. This selectivity was highly plastic and was sensitive to changes in the magnitude of rewards delivered at the end of the maze arms. End-arm selectivity was not dependent on the particular actions taken by the animal, nor was it directly linked to differences in the animals' runspeed; rather the ramping selectivity appeared to be linked to external features of the animals' environment. Finally, by performing concurrent recordings of dopamine concentration and local field potentials, we observed that beta band oscillations near goal reaching were inversely related to the spatial selectivity of the simultaneously recorded dopamine ramping indicating that dopamine signaling may influence synchrony in local striatal networks. We hypothesize that these prolonged, ramping dopamine signals may serve to motivate and direct spatial navigation behavior towards distant goals.

3.3.1 Sources of dopamine ramping

3.3.1.1 Uncertainty

We did not perform concurrent electrophysiological recordings of dopamine neuron firing in the midbrain, and therefore were not able to determine whether the ramping dopamine is a direct reflection of dopamine neuron firing patterns. Similar prolonged ramping changes in dopamine firing have been observed, rarely, in primate classical conditioning experiments (Fiorillo, Tobler et al. 2003; Fiorillo, Tobler et al. 2005). These signals were present during the delay period between the presentation of a conditioned stimulus indicating reward probability, and the upcoming reward delivery. The ramping firing rates were strongest following the conditioned stimulus associated with the highest uncertainty (50%) of upcoming reward delivery. Moreover, ramping was also observed when reward was delivered with 100% probability but reward magnitude disparities were large: cue presentation was associated with delivery of either a large or small reward (50% chance of each). The ramping was not present when magnitude discrepancies were small (small vs small or large vs large). We observed ramping signals on the T-maze in which there was a less than 100% probability that animals received reward, and we also observed ramping in the M and S mazes when animals were receiving nearly 100% reward but received different reward amounts in the two maze environments. It is possible therefore that a common mechanistic principle underlies our ramping signals and those observed by Fiorillo. I favor the explanation that the ramping signal is related to the magnitude of phasic dopamine received at the end of the maze, which is larger both in differential value conditions and in differential probability conditions (see model section below for explanation). Learning and online control of predictions and action choices is only necessary when environments contain some degree of unpredictability. When environments are entirely predictable, fixed action patterns can be implemented and no additional learning is needed.

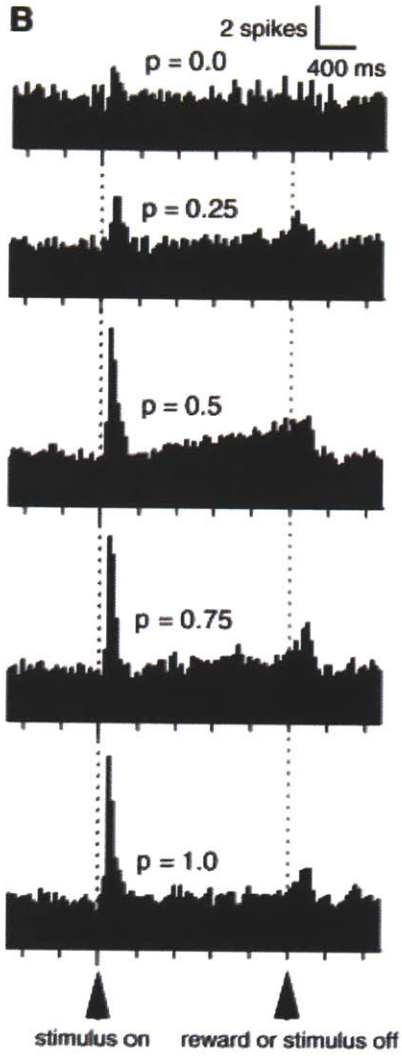

Figure 3.21 Ramping firing rate in dopamine neurons is related to reward uncertainty. Firing rate histograms for a putative dopamine neuron in the VTA recorded from a primate performing a probabilistic classical conditioning task. Each histogram represents the response of this neuron to stimuli associated with different reward probability. Ramping is observed between the stimulus and reward only in the 0.5 probability case. Adapted from (Fiorillo, Tobler et al. 2003).

3.3.1.2 Time, space, and path integration

The incremental increases in dopamine we observed indicate that they may be generated, in part, by a substrate (either internal or external) that changes incrementally during maze navigation. One possibility is that the ramping inversely reflects the estimated time to receive the reward. Timing signals have been observed in multiple brain areas, including the striatum, in rodents and primates (Buhusi and Meck 2005; Jin, Fujii et al. 2009). Presumed timing signals have also been shown to be weighted towards the time of estimated reward delivery (i.e. ramping to reward), and can carry information about preceding stimuli, perhaps in the service of working memory during simple instrumental conditioning tasks in primates (Gold and Shadlen 2007).

Another possibility is that dopamine ramping is an inversely weighted representation of distance remaining to reward reaching. Neural signals related to egocentric position have been measured in the firing of "grid cells" in the entorhinal cortex and hippocampal formation (Wilson and McNaughton 1993; Sargolini, Fyhn et al. 2006; Derdikman and Moser 2010). These signals could be used, in principle, to compute continuously updated distance estimates relative to rewards (i.e. ramping).

Representations of allocentric position or spatial context have been long known to be transmitted by the firing of "place cell" ensembles in the rodent hippocampus (Wilson and McNaughton 1993). Place cells fire at particular locations in an animal's environment during navigation, and place cells corresponding to reward locations are sometimes overrepresented (Hollup, Molden et al. 2001; Hok, Save et al. 2005; Hok, Lenck-Santini et al. 2007). These weighted contextual

representations could be used to generate the spatially selective ramping dopamine signals observed here (see model below).

3.3.1.3 Progressive accumulation and multiple transients

In theory, sustained elevated firing of dopamine neurons, not ramping, could result in progressive accumulation of dopamine in the extrasynaptic space, producing a ramping elevation in dopamine release until spillover rates and diffusion reach equilibrium with dopamine re-uptake (Arbuthnott and Wickens 2007). A potential problem for this hypothesis is the observation that the ramping signals scale to nearly equivalent peak magnitudes in the short and long mazes. Progressive accumulation of a constant elevated dopamine signal would presumably be larger in the long maze than the shorter maze, resulting in a larger peak. A constant sustained signal may be scaled also with maze length, however, so this possibility cannot be ruled out.

Another alternative is that multiple overlapping phasic transients could generate incrementally rising levels of dopamine that merge to produce what appears to be a steadily ramping signal. We did observe occasional discontinuities in the ramping signal which appeared to be linked to distinct events in the maze, such as the warning click to start the trial. In other instances, however, no such discontinuities were observed, and the ramping smoothly increased from the beginning of the maze to the goal reaching, making the possibility of superimposed transients less likely. There seems to be variability in the timecourse of "phasic" transients to cues and rewards observed here and in other studies (Day, Roitman et al. 2007; Gan, Walton et al. 2010). Primary reward evoked dopamine release can take 5-10 seconds to decay back to baseline levels,

whereas conditioned cue evoked release appears to be more transient on the range of 2-5s. This would provide ample time for integration of multiple overlapping transients during the course of a T-maze trial (5-8s). The typical rise time to peak of phasic transients is rapid (~1s), however, which is inconsistent with the generally slow progression of dopamine release observed in the T-maze. Regardless, the possibility of a summation of multiple continuous phasic release events cannot be excluded.

3.3.1.4 Pre-synaptic modulation

It is possible that our ramping dopamine signals do not originate directly from ramping firing in individual dopamine neurons. Dopamine release in the striatum is under the influence of pre-synaptic modulation from a variety of sources, both intra and extra striatal (Krebs, Trovero et al. 1991; Zhang and Sulzer 2003; Calabresi and Di Filippo 2008; Exley, McIntosh et al. 2012). Glutamatergic inputs from cortical and thalamic sources can facilitate the release of dopamine pre-synaptically through NMDA and mGluR receptors (Krebs, Trovero et al. 1991; Zhang and Sulzer 2003). Local cholinergic interneurons, which are themselves under control of cortical and thalamic inputs, can also extert a strong influence on dopamine release (Calabresi and Di Filippo 2008; Exley, McIntosh et al. 2012). Ramping firing rates have been observed for projection neurons in upstream structures such as the prefrontal cortex (Hok, Save et al. 2005).

3.3.1.5 Source of maze arm selectivity

The end-arm selectivity of the ramping dopamine signals indicates that features of the external environment must contribute. In principle, any set of features detectable by sensory systems,

such as aspects of the visual, auditory, or olfactory fields, could provide dopamine neurons with the representational substrate for value prediction. The possibility favored here is that the value prediction is not computed by inputs from any particular low level sensory system but is a product of a map of allocentric spatial context, such as that represented by the firing of hippocampal place cell ensembles (see model and above). Another possibility is that single isolated sensory cues can alter internally generated reward predictions, which are translated into proximity representations (ramping) by time or distance estimations. The sensory cues in the room are not weighted in any obvious way towards either side of the maze, so it is unclear why spatial selectivity in the ramping dopamine signals is biased towards the right side of the maze when rewards are equivalent on the two sides.

3.3.2 Dopamine signal heterogeneity

Electrophysiological studies of dopamine neuron activity have provided conflicting accounts of the uniformity of dopamine signaling (Fig 3.22). Some evidence suggests that dopamine signals are widely broadcast to target areas, while other studies suggest a degree of functional heterogeneity among dopamine neuron populations (Morris, Arkadir et al. 2004; Wightman, Heien et al. 2007; Matsumoto and Hikosaka 2009; Brown, McCutcheon et al. 2011). FCSV studies of dopamine release have widely found evidence in support of heterogeneous dopamine release, though the extent to which release in different striatal zones can show different functional properties is still not well understood. In this study, we have found evidence for both local homogeneity and also sparse functional heterogeneity. As a whole, the spatial selectivity of the signals across recording sites was highly correlated, both within and across subregions.

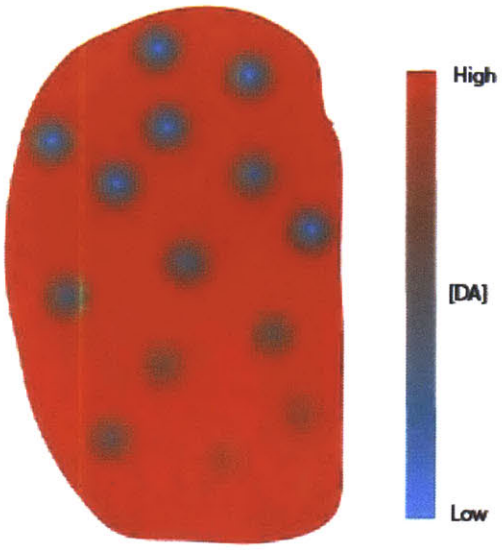

Figure 3.22. Model for observed patterns of dopamine release in different striatal regions in the maze task. The image shows a coronal section through the rat striatum and colors indicate relative levels of evoked dopamine release. The blue areas represent zones with dopamine suppression. The DLS has more regions with strong suppression than the VMS, which is relatively homogeneous in its release patterns.

Notably, however, we found a small subset of recording sites only in the dorsolateral striatum that showed negative ramping (progressive decline to goal), despite there being positive ramping in simultaneous recordings from other sites in the ventral striatum (Fig 3.7 and 3.8). This negative ramping was robust and consistent across sessions. We also observed rare instances of opposite spatial preferences in the ramping signals measured simultaneously in the DLS and VMS (Fig 3.14). These findings suggest that, while a large degree of functional homogeneity exists both within and across striatal subregions, some areas in the DLS exhibit signaling that is opposite to that in the VMS.

It is evident from this and other studies that the overall modulation of dopamine signaling in the DLS is weaker, and in this study, more heterogeneous (Zhang, Doyon et al. 2009; Brown, McCutcheon et al. 2011). It is not clear whether this is the result of active suppression of dopamine neuron firing or some unknown post-synaptic mechanism, such as re-uptake heterogeneity (Wu, Reith et al. 2001). Inhibitory inputs to DLS projecting SNc neurons arise from striosome neurons in the dorsal striatum, and according to some anatomical studies, from the ventral striatum as well (Haber, Fudge et al. 2000; Joel and Weiner 2000). Thus, it is

possible that ramping firing in VMS areas could actually generate suppression of DLS dopamine release (see model below). The explanation favored here is that DLS projecting dopamine neurons undergo mixed suppression and excitation during the maze task, creating a more heterogeneous release profile than that observed in VMS (Fig 3.22).

3.3.3 Impact of ramping dopamine signals on striatal networks

We have shown that a negative correlation exists between the spatial selectivity of the ramping dopamine signals and the strength of oscillatory power in the beta frequency range after goal reaching. This correlation does not appear to be present prior to goal reaching, indicating that a lack of dopamine doesn't by itself generate bursts of beta oscillations. I favor the hypothesis that dopamine in the striatum counteracts the natural processes that generate widespread synchronization of the network. In the case of the ventral striatum, we have shown that prominent global synchronization occurs most strongly in the beta range (see above), but in general dopamine may effectively counteract global synchrony in any frequency range that the network naturally supports. The function of dopamine in shifting network activity from a state of global low frequency synchrony to a spatially heterogeneous, high frequency regime has been suggested by pharmacological studies and in the context of pathologies of the dopamine system, such as Parkinson's disease, in which abnormal synchrony is observed in the absence of dopamine (Berke 2009; Jenkinson and Brown 2011). Our work suggests that fluctuations in dopamine signaling in the normal striatum during behavior may be important for determining network states during the learning and execution of goal directed behaviors. Higher frequency

local synchrony promoted by dopamine signaling may be ideal for enabling network plasticity and flexibility. Of course, our evidence for dopamine's role in controlling oscillatory state is correlational rather than causal. To test this idea directly, a manipulation of dopamine signal, for example with channelrhodopsin stimulation of dopamine neurons, is necessary (Witten, Steinberg et al. 2011).

It is interesting to note that the beta oscillations appear most strongly after goal reaching, whereas the ramping peaks well before the beta power emerges (Fig 3.20). The ramping dopamine does differentiate the two end-arms after goal reaching, however. This could be because of delays in dopamine clearance or because the active process that generates the arm selectivity continues to influence release after goal reaching and reward consumption. In any case, the post-goal selectivity appears to influence neural activity dynamics in the ventral striatum.

We were not able to record enough single unit data to investigate the effect of dopamine ramping on striatal spiking. The influence of dopamine on excitability in the striatum is highly cell-type dependent, so it is possible that the ramping signal can both suppress and elevate firing probabilities in different subtypes (Surmeier, Ding et al. 2007). Moreover, the different dopamine receptor subtypes in the striatum (D1 and D2) have different affinities for dopamine meaning that the ramping dopamine may activate the higher affinity receptor (D2) more strongly (Richfield, Penney et al. 1989).

3.3.4 Behavioral functions of the ramping dopamine signals

The current experiments were not designed to test whether the novel ramping dopamine signals are necessary or sufficient for different aspects of learning, motivation, and spatial navigation. The prolonged timecourse of the dopamine signaling indicates that it may be ideal for modulating aspects of ongoing behavior, such as motivation or effort to pursue goals. Dopamine's involvement in maintaining motivation has been highlighted in the profound motivational deficits present in Parkinson's disease. Pharmacological manipulations in rats have also indicated that dopamine signaling through striatal D2 receptors is critical for maintaining effort to reach distant goals in maze tasks (Whishaw and Dunnett 1985; Salamone and Correa 2002). Other studies have proposed a role for dopamine in mediating "incentive salience" or motivational responses to environmental cues (Berridge 2007). These ideas are consistent with a potential role of our spatially selective ramping dopamine signal in driving locomotion toward goals in a context dependent manner.

3.3.4.1 Learning in uncertain environments

A strong focus in the literature has been on dopamine's role in driving learning by providing feedback signals in response to unexpected rewards (Schultz, Dayan et al. 1997; Schultz 2002). Although a direct causal role in the reinforcement learning process has been disputed (Darvas and Palmiter 2009), it is clear, at least at a molecular level, that dopamine signals are capable of promoting long lasting functional changes in striatal circuits via LTP and LTD (Reynolds and Wickens 2002). Here, we have observed phasic dopamine signaling to cues and unexpected rewards which closely parallel those proposed in other studies to be important for driving

plasticity and learning (Schultz, Dayan et al. 1997; Day, Roitman et al. 2007; Wickens, Horvitz et al. 2007). Reinforcement learning theory proposes that these phasic dopamine transients are used to "stamp in" neural representations of actions and predictive sensory input that precede unexpected rewarding events (Schultz 2002; Wickens, Horvitz et al. 2007). This information can then be used to guide animals' subsequent behavior. For this "stamping in" process to work, traces of neural activity that represent the preceding actions or cues to be reinforced must be present at the time of the phasic dopamine reward transient, a computational challenge referred to as the "temporal credit assignment problem" (Sutton 1998). Several possible solutions to this problem have been proposed, such as replay reactivation and neural eligibility traces (Louie and Wilson 2001; Foster and Wilson 2006; Walsh and Anderson 2011). Another potential solution to this problem is to use a reward predictive signal that is present prior to reaching the actual reward, to pre-emptively learn about the actions taken to reach the reward. The predictive signal itself would be anchored to an internal or externally derived substrate, such as time, distance, or spatial location (see above). The ramping signals described here fit the criteria for such a signal and could in theory be used to update circuits based upon ongoing reward predictions. This proposal is consistent with the idea that these signals are strongest when multiple, alternative courses of action are present in a given environment.

3.3.4.2 Spatial navigation

To our knowledge, ramping dopamine signals in the striatum have not been observed in the standard operant or classical conditioning tasks studied with FSCV to this point. It is possible that these signals play a special role in directing motivation along particular,

learned routes through space. The importance of dopamine in spatial navigation has been highlighted by a number of behavioral studies (Whishaw and Dunnett 1985; Salamone, Correa et al. 2007; Braun, Graham et al. 2012). Modeling work by Foster and Dayan has proposed that phasic dopamine signals to unexpected goals in a water maze could be used to create spatial value gradients that could be used to guide navigation (Fig. 3.23; (Foster, Morris et al. 2000). These value gradients closely resemble the ramping dopamine signals we have observed in our study. I propose here that dopamine ramping signals

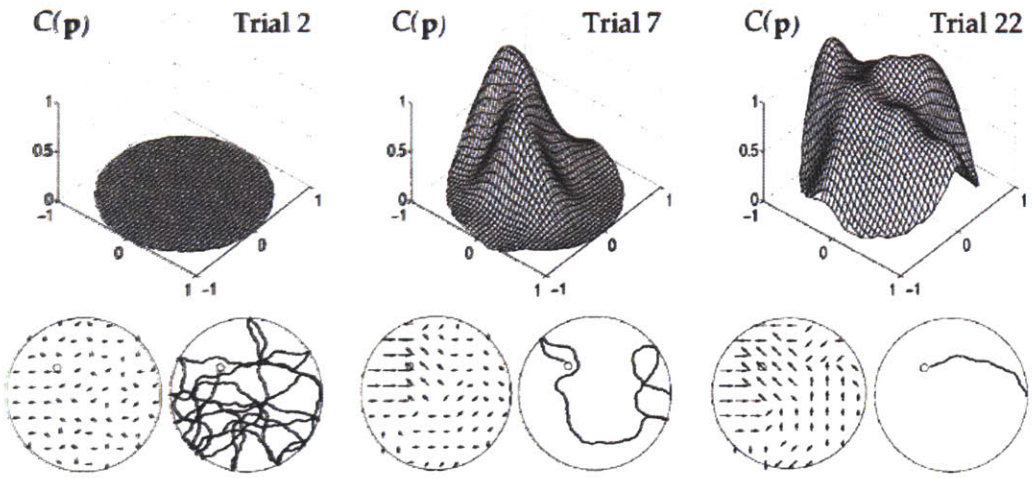

Figure 3.23. Foster and Dayan model for learning of place cell generated value functions. This model was designed to explain how rats might use spatial representations to choose the optimal sequences of actions to find a platform in a Morris Water Maze using an actor/critic neural architecture. For each trial, the critic's value function $C(p)$ is shown in the upper, three-dimensional plot; at lower left, the preferred actions at various locations are shown ; at lower right is a sample path. Trial 2: After a timed-out first trial, the critic's value function remains zero everywhere, the actions point randomly in different directions, and a long and tortuous path is taken to the platform. Trial 7: The critic's value function having peaked in the northeast quadrant of the pool, the preferred actions are correct for locations close to the platform, but not for locations further away. Trial 22: The critic's value function has spread across the whole pool and the preferred actions are close to correct in most locations, and so the actor takes a direct route to the platform. Adapted from Foster and Dayan 2000.

participate directly in guiding spatial navigational behavior. As animals move through an environment, the dopamine signal is updated based on the predicted "value" of moving from one location to the next in the spatial sequence along a route to a predicted goal. As long as the dopamine signal continues to rise (or is maintained) with each successive step, the animal continues to exhibit goal-directed locomotion. If the animal suddenly enters a location that is not a member of a valuable, goal-directed route, the behavior becomes flexible to change, and exploratory or stopping behavior may be initiated. This theory also implies that locomotion behavior in environments that are close to predicted rewards is more robust against external distractions than behaviors that are far from predicted rewards. This type of motivation based upon value predictions and reward proximity may not be limited to actions performed in space, but may also extend to prolonged motivated behaviors in time.

3.3.5 Model for generation of ramping dopamine for value based spatial navigation

Here I will present a simple qualitative model for how the ramping dopamine signals may be generated through repeated experience with unexpected rewards or goals. This is only one of many possible mechanisms of course, and further experiments will be needed to test the contribution of different brain areas to the ramping dopamine signals. The model should be evaluated mainly on the efficiency and plausibility of the principles presented, as the exact neural circuits involved are unknown.

3.3.5.1 Background

The main concept behind this model is that the brain utilizes representations of external, allocentric space to create "motivational gradients" that signal the relative values of spatial locations, in the form of the ramping dopamine signals, to drive goal seeking locomotion. The basic substrate (as in the model proposed by Foster and Dayan (Foster, Morris et al. 2000)) is the spatial map represented in the ensemble firing of hippocampal place cells (for example, pyramidal cells in the CA1 subregion), which fire when an animal is in a particular location in a familiar environment. The spatial location information is transmitted through the hippocampal output structure, the Subiculum, to a range of downstream targets, including areas of the ventral medial striatum (Floresco, Todd et al. 2001; Lisman and Grace 2005). Neurons of the ventral striatum project to dopamine neurons both indirectly, via the ventral pallidum, and directly, providing potential means by which hippocampal spatial maps could influence dopamine neuron firing (Fig 3.24). This pathway has been shown to be important for controlling population firing of dopamine neurons, which is believed to be related to tonic dopamine release in the striatum (Floresco, West et al. 2003). A direct projection from area CA3 to the VTA has also been discovered, providing a direct means of interaction (Luo, Tahsili-Fahadan et al. 2011). Dopamine innervation is stronger to the striatum than to any other structure in the brain, including the hippocampus, creating a potential feedback loop between the ventral striatum projection neurons and dopaminergic neurons of the midbrain.

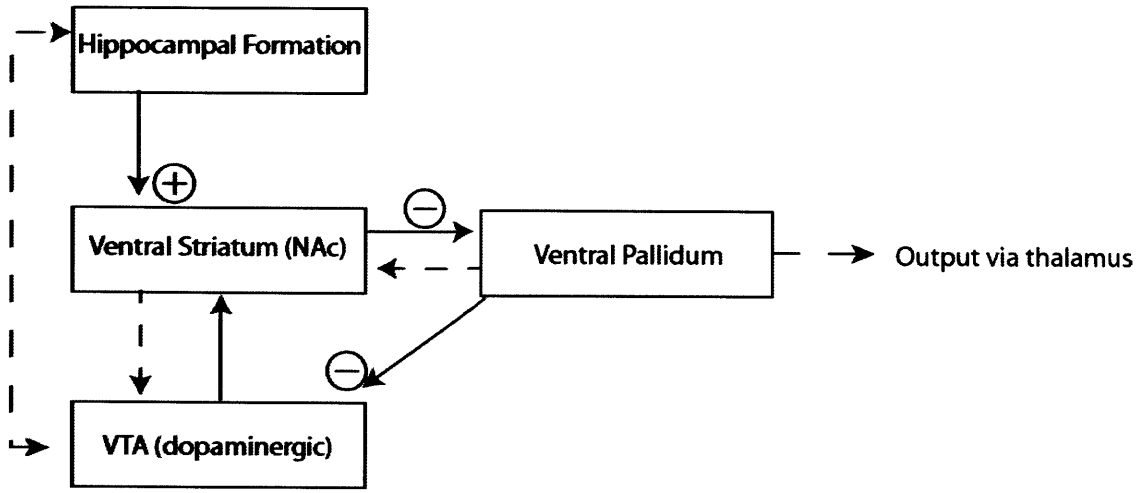

Figure 3.24 Simplified schematic of relevant connectivity for the proposed model. Solid lines represent connections that are critical for the proposed model, while dashed lines represent alternative routes through which the model could be implemented.

One clue that an interaction between hippocampus and ventral striatum may be critical for generating the ramping dopamine responses comes from electrophysiological recordings in the ventral striatum in maze tasks. Work both in our lab on the previously described T-maze task (Fig 3.25A and B) and from a number of other labs has shown that firing rates of medium spiny projection neurons slowly ramp up to goal reaching, much in the same manner that our dopamine signal does (Hikosaka, Sakamoto et al. 1989; Lavoie and Mizumori 1994; van der Meer and Redish 2011). Moreover, work in the Redish lab has found that ramping cells in the ventral striatum display phase precession relative to hippocampal theta rhythms (van der Meer and Redish 2011). Puzzlingly, despite this and other electrophysiological evidence of hippocampal/striatal interactions (DeCoteau, Thorn et al. 2007; Tort, Kramer et al. 2008; Lansink, Goltstein et al. 2009) there have been conflicting reports of spatial coding in the ventral striatum (Eschenko and Mizumori 2007; Berke, Breck et al. 2009; Lansink, Jackson et al. 2012).

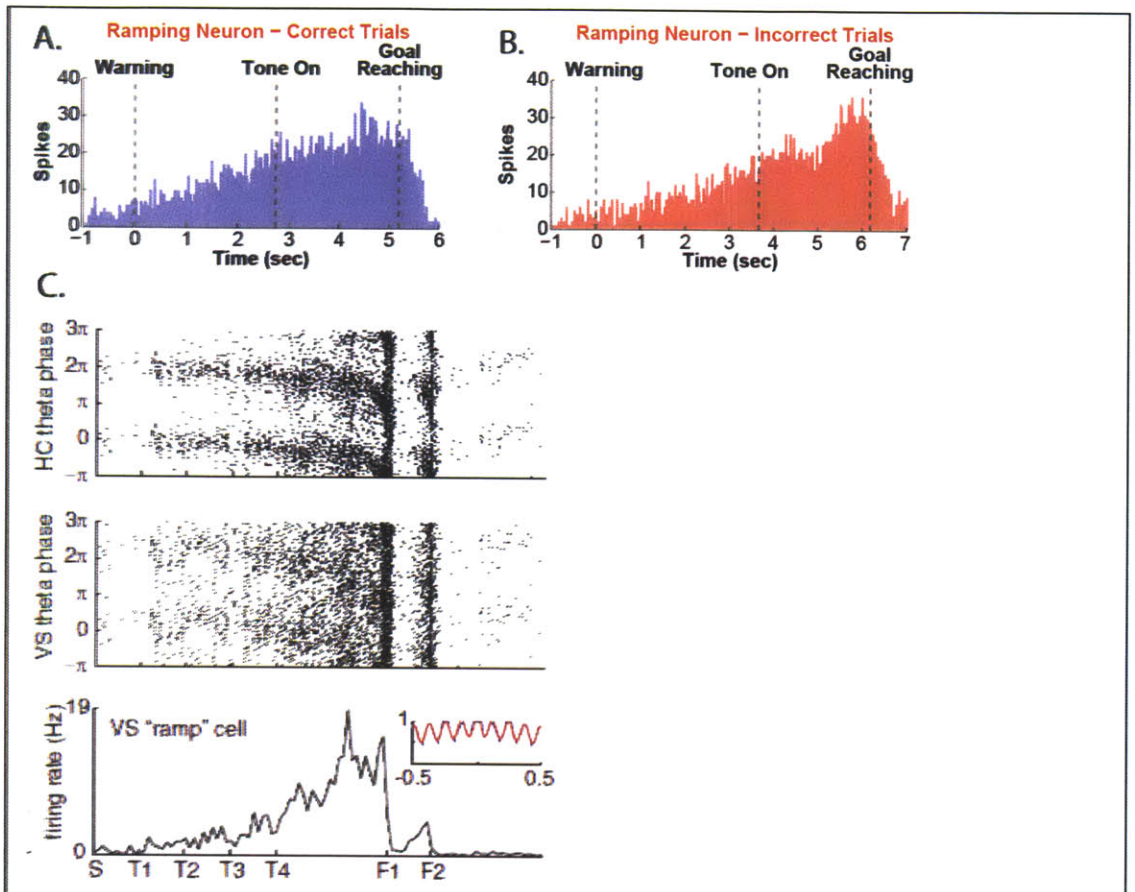

Figure 3.25 Ventral striatum neurons show ramping firing rates during maze tasks and are coordinated with hippocampal theta rhythms. A and B. Average firing rate histograms from a ramping neuron recorded in ventral striatum during correct (A) and incorrect (B) trials in a single 40 trial session on our T-maze task. From Atallah et al. in prep **C.** A ramping neuron in the ventral striatum that undergoes phase precession with respect to hippocampal and ventral striatal theta rhythms. The x axis represents spatial position in the maze task. From (van der Meer and Redish 2011)

Pioneering electrophysiological recordings in the hippocampus in maze navigation tasks have revealed that place cells display temporally compressed "replay" during sleep and at reward consumption periods during active maze running (Louie and Wilson 2001; Foster and Wilson 2006). During these events, place cells that were active during the preceding maze runs are sequentially reactivated in reverse temporal order during sharp-wave ripple events that occur

when an animal stops to consume reward. Some reactivation has also been observed in the ventral striatum, though this does not seem to display the same tight temporal sequencing that hippocampal replay does (Lansink, Goltstein et al. 2009). Interestingly, this ventral striatal replay seems to be weighted towards the representations of reward sites.

The evidence highlighted above suggests that hippocampal, ventral striatal, and dopaminergic circuits may interact to assign relative predictive values to different locations in space. I believe that an end result of this interaction is the generation of spatially selective, ramping dopamine signals that we have measured in this study. Below is a mechanistic, circuit-level model for how these ramping signals may be generated with repeated experience with unexpected rewards after spatial navigation.

3.3.5.2 Model

This qualitative model is based on the anatomical connectivity patterns outlined above. Multiple hippocampal output neurons from the subiculum converge onto the dendritic trees of single medium spiny output neurons of the ventral striatum (Shepherd 2004). Subiculum neurons have spatially defined firing fields, similar to place cells in the hippocampus (Sharp and Green 1994). In the early, exploratory phase of behavior, the spatial environment is uniformly represented, so no subicular inputs are weighted more strongly than others (Fig 3.26). When the animal encounters an unexpected reward, several events occur. First, as previously highlighted, a phasic pulse of dopamine release (a "reward prediction error" signal) is generated in the midbrain dopamine neurons and this release is strongest in the ventral medial region of the striatum (Brown, McCutcheon et al. 2011). Second, a sharp-wave ripple replay event is triggered in place

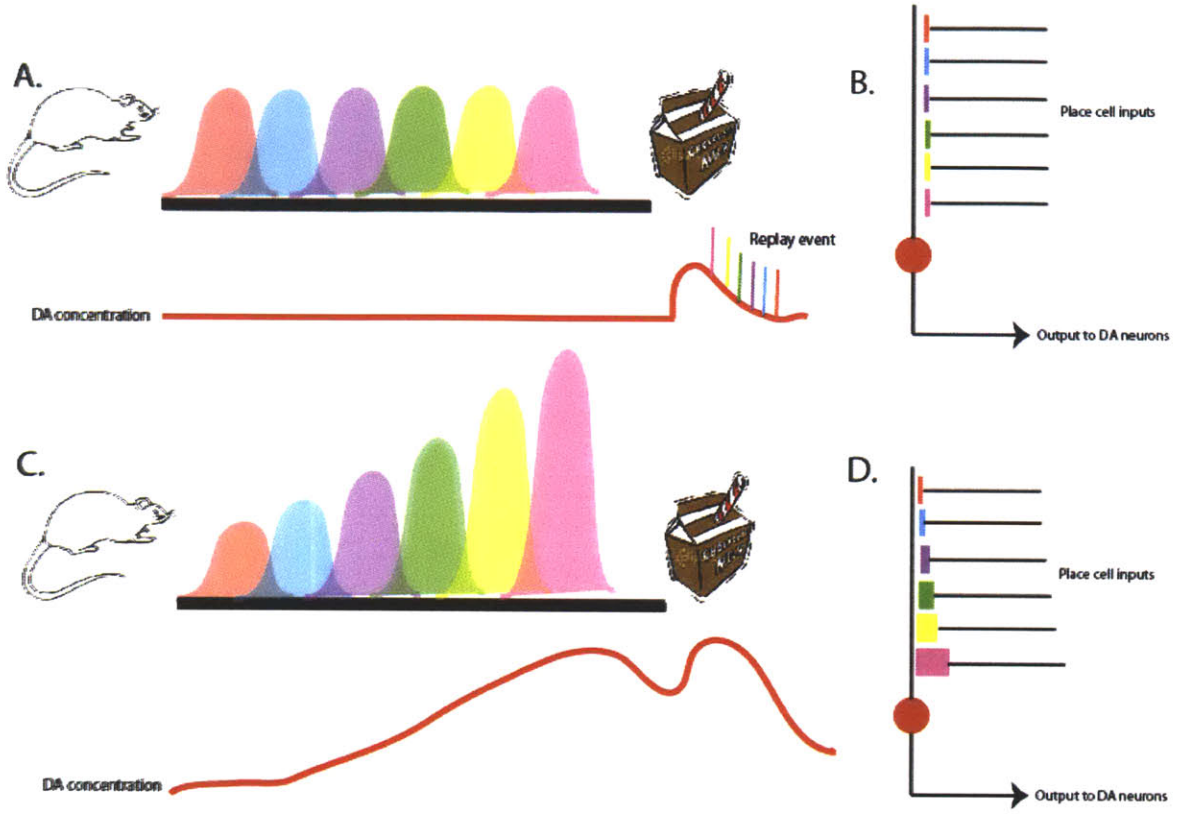

Figure 3.26. Schematic of the proposed model for generation of ramping dopamine signals via phasic dopamine feedback signals. A. Early exploratory phase. The rat navigates along a 2 dimensional trajectory, such as a T-maze track to find a chocolate reward. The spatial tuning curves of each place cell are equivalent at this stage (each color represents the weighted tuning curve input from each place cell), as are their synaptic inputs onto cells of the ventral striatum **(B)** (in principal this could be one of many different brain areas). The unexpected reward triggers a phasic pulse of dopamine in the striatum (red trace lower) which coincides with a reverse replay of the place cells that were active on the track. **C and D.** Goal-directed phase. The consequence of the events in A is that the spatial inputs onto the striatum become weighted through experience dependent plasticity towards the reward location. The output of the ventral striatal neurons is similarly skewed (ramping), which generates a ramping signal in the downstream dopamine projections.

cells of the hippocampus. Presumably, this translates to a parallel replay event in the downstream output neurons of the subiculum (though this has never been shown experimentally). The result of this process is that inputs onto single MSNs are sequentially activated in reverse temporal order. The effect of the coincident pulse of dopamine with the incoming sequential inputs is to asymmetrically strengthen synaptic weights from each of the subiculum cells onto the MSN through LTP (Reynolds and Wickens 2002). Dopamine signaling through D1 receptors is critical for LTP in the striatum and is strongly dependent on the timing of the dopamine pulse relative to post-synaptic dendritic depolarization (Reynolds and Wickens 2002; Shen, Flajolet et al. 2008). Inputs that arrive closer in time to the dopamine pulse are strengthened more effectively than those that arrive later. This creates a ramping response in the ventral striatum MSNs during subsequent traversals of the routes leading to goals. The ramping firing can be transmitted downstream (either directly or indirectly through the ventral pallidum) to dopamine neurons, which can produce the action motivating ramping signal that we have measured.

In environments with multiple goal sites, such as on a T-maze, the magnitude of the synaptic modification by the phasic dopamine pulses at goal reaching determines the strength of the predictive ramps leading to that goal. The size of phasic signaling at goals is dependent on at least 2 factors: prior predictions (reward probability) and relative magnitude or value (Schultz, Dayan et al. 1997; Tobler, Fiorillo et al. 2005). Less predictable rewarding events will lead to a larger dopamine "reward prediction error." Fully predicted, but more valuable reward options (relative to other rewards in the same context) will also elicit larger dopamine transients.

3.3.5.3 Predictions of the model, outstanding questions, and future experiments

One prediction of the proposed model is that dopamine ramping should be directly proportional to the relative magnitudes of the phasic dopamine signals at different goal locations. Routes to locations with uncertain reward probabilities and/or high magnitude rewards should elicit larger ramping signals relative to alternative trajectories. This latter prediction is supported by our data, while the former was not explicitly tested for in the current experimental paradigm.

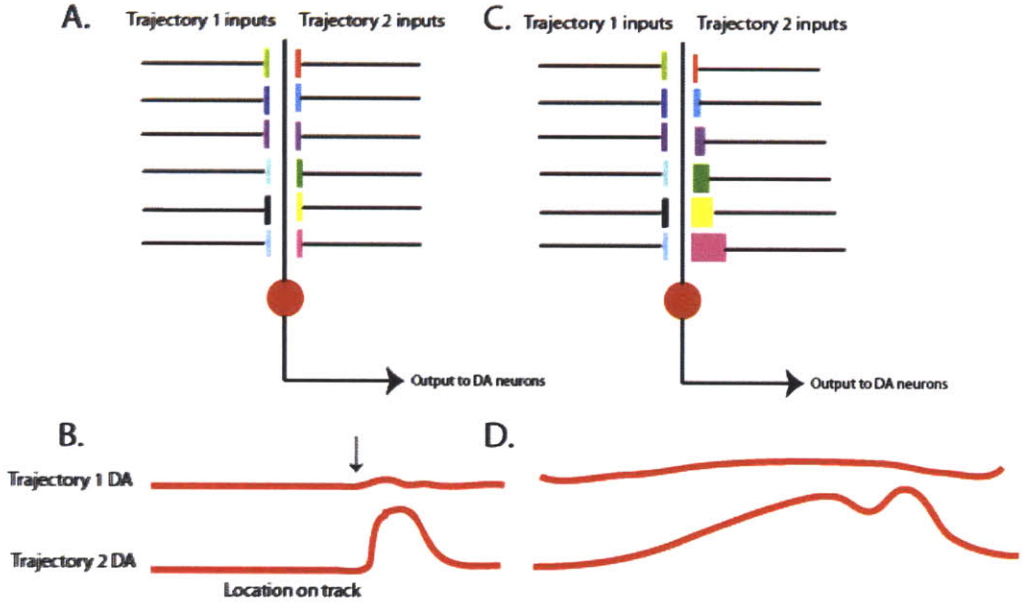

Figure 3.27 Generation of trajectory selective value signals. A. Early exploratory phase as in Fig 3.26 in a spatial environment with 2 distinct trajectories (such as a T-maze). The inputs converging onto the single ventral striatum neuron are from place cells that fire selectively in a single location in one of the two trajectories. **B.** As in Fig 3.26, the rat encounters unexpected reward at the ends of both trajectories (location indicated by arrow), but he receives a larger reward in trajectory 2, which results in a larger phasic dopamine signal at this location. **C.** Through the replay mechanism described in Fig 3.26, the inputs corresponding to locations in trajectory 2 are asymmetrically strengthened, while the inputs corresponding to trajectory 1 are not. **D.** This results in a larger ramping dopamine signal for trajectory 2 relative to trajectory 1.

Because we could not readily discriminate phasic transients from preceding ramping signals in the majority of cases, we could not directly test whether phasic signals were proportional to the ramping signals. However, dopamine signals continued to discriminate the two end-arms well after goal reaching, indicating that perhaps the superimposed phasic signals do relate to the preceding ramping signals (Fig. 3.21C). This prediction could be formally tested with direct electrophysiological recordings of dopamine neuron firing.

Another testable prediction of the model is that the dopamine ramping signals, or at least the spatial selectivity of these signals is dependent on intact spatial processing by the hippocampus. Inactivation of the hippocampus or subiculum, for example through optogenetic means, could test this prediction. Muscimol inactivation of the ventral `subiculum has been shown to reduce tonic population firing of dopamine neurons, indicating a possible link between the hippocampus, tonic dopamine neuron firing and the spatially selective ramping signals we observe (Floresco, West et al. 2003). Restricting optical inactivation to maze running would test the necessity of the hippocampal spatial map for maintaining the ramping and the spatial selectivity of the ramping signals, whereas optical inactivation during reward consumption could test the hypothesized role of hippocampal replay in forming the dopamine gradients. More selective manipulations could target terminal projections from the hippocampus to different regions (such as the ventral striatum, PFC, VTA, ect.) to test the precise circuit-level interactions involved in generating the ramping responses.

One way to test the hypothesis that "spatial values" are stored in the relative synaptic weights of subiculum/hippocampus neurons onto MSNs (or neurons of PFC, VTA, ect), would be to run an

experiment in which the values of spatial locations change flexibly from trial to trial in a predictable manner. This could be done by delivering cues that signal the value of the upcoming reward during navigations through the same maze route. If the ramping signals were tied entirely to the spatial map substrate, they would not depend on the cue value on a trial to trial basis. If instead, the ramping signals reflected a prediction that is independent of previously modified synaptic weights of place cells, the magnitude of the ramping would change rapidly on a trial to trial basis as the value cues do. Notably, this experiment would not definitively test the contribution of a spatial map substrate.

It remains unclear what the relationship (if any) is between the "uncertainty" related ramping signals in dopamine neurons observed by Schultz and Fiorillo and those described here. The model suggests a possible unification of these observations. Both the Fiorillo ramping and our ramping signals may be linked to continuously updating substrates: in the former case, time to reward, and in the latter case, location in space or proximity to goal. The ramping in both cases may be driven by asymmetric synaptic plasticity by dopamine transients, as described in the model. In the Fiorillo study, the ramping was largest for the uncertain reward case because the reward transients were most frequent in that condition. However, this view is difficult to reconcile with the motivational roles ascribed above to the dopamine signals, as it implies that motivation would be highest when pursuing uncertain goals. Given this view, it is perhaps more likely that the dopamine signals are in fact driving online learning of actions and stimuli (in either highly valued, or highly uncertain conditions) as discussed above, rather than directly motivating ongoing behaviors.

Finally, it remains an outstanding question what the primary effect of the dopamine ramping signals are on the striatal network. The model favored here proposes that the ramping firing rates observed in some cells of the ventral striatum are active in generating the dopamine responses rather than the converse. A testable prediction related to this proposal is that ramping neurons of the ventral striatum should display spatial selectivity in line with the dopamine selectivity, only when there are relative differences in reward related transients in an environment (i.e. when rewards differ in probability or value). This may account for the conflicting accounts of "spatial selectivity" in firing patterns of ventral striatal neurons. Ventral striatal neurons are perfectly poised to discriminate spatial trajectories based on their relative values in the service of directing action and motivation.

I favor this explanation because the immediate effects of dopamine are modulatory in nature and are unlikely to produce direct changes in firing rate along the timecourses we observe (Surmeier, Ding et al. 2007). More likely, the dopamine signals are broadcast to wide regions of the striatum (including the more dorsal areas) and have a complex effect on striatal output including, as noted here, reducing synchronous low frequency oscillations (Jenkinson and Brown 2011).

3.3.5.4 Implications

It has long been believed that the ventral striatum plays a crucial role in motivating and reinforcing behaviors leading to highly valued goals, but the neural mechanisms underlying this function have been unclear (Mogenson, Jones et al. 1980; Humphries and Prescott 2010). According to the proposed model, the ventral striatum may accomplish this, in part, by transforming representations of spatial context in the hippocampus into motivational dopamine

gradients that can direct locomotion to distant goals. The essence of the model is that experience with rewards generates predictions that are "smeared" across the spatial substrate creating what can be thought of as a weighted map of space. How these dopamine signals are used to direct motivation, navigation, and learning remain unclear, but the possible implications are far-reaching. Goal-directed behaviors are ubiquitous among all mammalian species, including humans, and predictions based on prior experience are necessary for proper execution and learning of these behaviors. Predictions, in the form of ramping ongoing dopamine signaling, may provide the drive for motor output systems to invigorate movement towards distant goals.

In addition to the insights into mechanisms of normal goal-directed behaviors, this finding could also hold relevance for understanding dysfunctions of the dopamine system. For example, this finding may help to explain the uncontrollable drive to seek drugs that addicts experience in contexts that they've previously used in. Artificially elevated dopamine levels by drugs of abuse may tag the representations of the spatial context with motivational value. According to the proposed model, this would result in prolonged elevations in dopamine signaling in the ventral striatum upon subsequent exposure to the context, which could drive drug seeking. Several groups have proposed models for drug abuse in which artificial transients mimic naturally occurring prediction errors to tag cues with motivational salience (Redish, Jensen et al. 2008; Schultz 2011). The difference with the model proposed here is that the weighted tagging of spatial context would generate ongoing, prolonged dopamine signals, while these other models predict only isolated, cue-evoked transients, which in my view are unlikely to be sufficient to produce the ongoing drives experienced by drug users. If spatial substrates from the

hippocampus are involved in the drug seeking process as hypothesized, this could suggest new options for intervention in severe drug users.

The model I propose suggests that spatial maps in the hippocampus are the central contributor to the ramping dopamine signals and their spatial selectivity, but as previously discussed, other factors may be involved. Proximity to goals may be represented in a variety of internal and external reference frames including: time, distance, and landmark tracking (multiple "signpost" cues). Any or all of these factors might be used by the brain to construct predictions about distant rewards. The striatum is capable of representing all of these substrates in various forms, so it is likely that the "tagging" principles outlined here can be generally applied to attach motivational value to these substrates. Indeed, it may be that this is a central function of the loops through the ventral striatum and downstream basal ganglia structures: to transform positive experiences into predictions that can drive action.

Classic studies of dopamine signaling have highlighted the importance of cue-evoked phasic transients for motivating behaviors (Berridge 2007; Flagel, Clark et al. 2011; Schultz 2011). I would argue that these studies largely model the highly artificial case in which a single modality cue holds the majority of the predictive information about upcoming rewards (i.e. classical conditioning). The ramping signals measured here may represent the more general case in which the predictive substrate (time, space, distance) is multimodal and changes incrementally in a predictable manner during goal-directed behavior.

3.4 Methods

Headstage implantation Male Long Evans rats (n = 9) were deeply anaesthetized and, with sterile precaution were implanted with headstages carrying two independently movable voltammetry microsensor probes targeting the right (n = 3), left (n = 4), or bilateral (n = 2) DLS (AP +0.5 mm, ML ±3.5 mm, DV 3.5-4.0 mm), two probes targeting the VMS of the same hemisphere (AP +1.5 mm, ML ±2.1 mm, DV 6-7 mm), and a Ag/AgCl reference electrode in the contralateral posterior cortex (AP -2.3 mm, ML ±3.5mm, DV ~0.5mm) according to approved surgical procedures (Barnes, Kubota et al. 2005). Four rats carried 8-16 independently movable tetrodes, in addition to the voltammetry probes, for simultaneous neuronal recordings. Five rats were implanted with tungsten bipolar stimulation electrodes (FHC Inc.) straddling the ipsilateral medial forebrain bundle (AP -4.6 mm, ML ±1.3 mm, DV 7-8 mm) to verify striatal dopamine release (see below).

Behavioral training All behavioral training was conducted on a custom built "grid maze" with fully reconfigurable tracks and walls. Approximately 4 weeks after implantation, training began on a conditional T-maze task with auditory instruction cues (Barnes, Kubota et al. 2005). Daily behavioral sessions consisted of 40 trials. Each trial began with a brief warning click, followed 0.5 s later by the lowering of a swinging gate, allowing the rat to run down the maze. Half-way down the long arm, a tone was triggered (1 or 8 kHz), indicating which end-arm to visit in order to retrieve chocolate milk reward delivered through automated syringe pumps (Syringe Pumps Inc.) upon arrival. The spatial position of the rat was continually monitored via video tracking (Neuralynx Inc.). Tone delivery and syringe pumps were controlled by in-house behavioral

software written in Matlab (Mathworks Inc.). After 15-35 T-maze sessions, a group of rats (n = 3) received 17 training sessions on the "M-maze" task in which the end-arms of the T-maze were extended. These rats received a larger amount of reward (0.4 ml) at the goal site in one end-arm than at the other (0.1 ml). After 2-3 sessions with a given set of spatial reward contingencies, the reward amounts in the two end-arms were reversed. One rat was required to make turn choices in response to tones as in the previous T-maze task, whereas the other two rats were directed randomly to one end-arm of the maze on each trial by removing the track to the opposite arm (20 trials to each arm) without tone presentation. Another group of rats (n = 2) was trained on the S-maze task. These rats were simply required to run back and forth to retrieve a large volume of chocolate milk (0.4 ml) at one end of the maze and a small volume (0.1 ml) at the other end. Consecutive visits to the same reward site did not trigger reward pumps.

Voltammetry data acquisition and analysis_Waveform generation and data acquisition for voltammetry recordings were carried out by two PCI data acquisition cards and software written in LabVIEW (National Instruments). Triangular voltage waveforms were applied to chronically implanted carbon fiber electrodes, relative to the reference electrode, at 10 Hz. Electrodes were held at -0.4 V between scans, and ramped to 1.3 V and back to -0.4 V during each scan (Clark, Sandberg et al. 2010). Current produced by redox reactions during voltage scans was recorded.

We compiled a library of current vs. applied voltage (C/V) templates for dopamine and pH changes of various magnitudes by stimulating the medial forebrain bundle to induce dopamine release in the striatum in 5 rats under isofluorane anesthesia. We used these templates to perform chemometric analysis (Keithley, Heien et al. 2009) on voltammetry measurements obtained during behavior using in-house MATLAB software. This procedure allowed us to distinguish

changes in current due to dopamine release from changes due to pH or to other electroactive substances. Current changes were converted to estimated dopamine concentration by using calibration factors obtained from *in-vitro* measurements of fixed dopamine concentrations. Behavioral video tracking were synchronized with voltammetry recordings by marker TTL signals sent to the voltammetry data acquisition system.

Average traces were computed by first normalizing dopamine concentrations to the peak values that occurred during each session, averaged across all trials of the session, for each individual probe. These normalized recordings were then averaged across probes and sessions. Maze arm selectivity was computed by the following equation:

$$Dopamine\ Selectivity\ Index = \frac{\overline{[DA]}_{Left\ Arm} - \overline{[DA]}_{Right\ Arm}}{\overline{[DA]}_{Left\ Arm} + \overline{[DA]}_{Right\ Arm}}$$

Extended increases and decreases in dopamine were identified by analyzing trial-by-trial recordings in a 0.5-s window prior to goal reaching. T-tests ($P < 0.05$) were conducted on these trial averages to determine whether responses showed a significant increase or decrease prior to goal-reaching. Responses were not selected explicitly for ramping profiles. Trials with significant responses to the warning click were identified by analyzing average dopamine concentrations within a 1-s time-window following warning click. Trials with isolated goal-related transients were those that showed a significant (t-test, $P < 0.05$) increase in dopamine concentration after goal-reaching but did not show a significant increase in the 1-s window prior to goal-reaching. Population selectivity was assessed for significance by comparing the variance of selectivity indices from the T-maze recordings to the set of variances obtained by shuffling the dopamine concentrations on the two end-arms and bootstrapping 10,000 times. To identify

changes in selectivity and ramping magnitude across training, Z-scores of selectivity indices and response magnitudes were computed by normalizing average dopamine values in a 1-s time-window prior to goal-reaching across sessions for each rat. Choice selectivity was computed similarly to the selectivity score for dopamine:

$$Behavioral\ Selectivity\ Index = \frac{Left\ Arm\ Choices - Right\ Arm\ Choices}{Left\ Arm\ Choices + Right\ Arm\ Choices}$$

Neuronal data acquisition and analysis_Local field potentials (LFPs) were recorded simultaneously with dopamine in the VMS throughout training in 4 rats (Howe, Atallah et al. 2011). LFP activity was amplified and sampled with a digital acquisition system (Neuralynx Inc.). All LFPs were referenced relative to a local tetrode placed in the striatum 2-3 mm dorsal to the recording sites. Spectral analysis was performed with in-house MATLAB software and with the chronux toolbox (http://chronux.org). Electrical artifacts produced by voltammetry scanning were removed by bandpass filtering the signal at 10 Hz. Spectrograms were constructed by the multitaper method, with three tapers, a time bandwidth product of 2, and a window width of 0.75 s. Selectivity indices for 15-28 Hz beta power were computed similarly to those for dopamine. Mean beta power values for modulation indices were averaged within a time-window after goal-reaching in which beta power was maximal (0-1 s post-goal-reaching).

Histology Tetrode and probe positions were verified histologically (Barnes, Kubota et al. 2005). Brains were fixed by transcardial perfusion with 4% paraformaldehyde in 0.1 M NaKPO4 buffer, post-fixed, washed in the buffer solution, and cut transversely at 30-μm on a freezing microtome, and they were stained with cresylecht violet to allow reconstruction of the recording sites. For a

subset of the voltammetry microsensor probes, a constant current (20 mA, 20 s) was passed through the probe prior to fixation to produce a micro-lesion to serve as a landmark for probe-tip locations.

References

Alexander, G. E., M. R. DeLong, et al. (1986). "Parallel organization of functionally segregated circuits linking basal ganglia and cortex." Annu Rev Neurosci **9**: 357-381.

Arbuthnott, G. W. and J. Wickens (2007). "Space, time and dopamine." Trends Neurosci **30**(2): 62-69.

Atallah, H. E., D. Lopez-Paniagua, et al. (2007). "Separate neural substrates for skill learning and performance in the ventral and dorsal striatum." Nat Neurosci **10**(1): 126-131.

Baker, S. N., E. Olivier, et al. (1997). "Coherent oscillations in monkey motor cortex and hand muscle EMG show task-dependent modulation." J Physiol **501 (Pt 1)**: 225-241.

Balleine, B. and A. Dickinson (1992). "Signalling and incentive processes in instrumental reinforcer devaluation." Q J Exp Psychol B **45**(4): 285-301.

Balleine, B. W. and A. Dickinson (1998). "Goal-directed instrumental action: contingency and incentive learning and their cortical substrates." Neuropharmacology **37**(4-5): 407-419.

Barnes, T., Y. Kubota, et al. (2005). "Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories." Nature **437**: 1158-1161.

Barnes, T. D., Y. Kubota, et al. (2005). "Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories." Nature **437**(7062): 1158-1161.

Benchenane, K., A. Peyrache, et al. (2010). "Coherent theta oscillations and reorganization of spike timing in the hippocampal- prefrontal network upon learning." Neuron **66**(6): 921-936.

Berke, J. D. (2008). "Uncoordinated firing rate changes of striatal fast-spiking interneurons during behavioral task performance." J Neurosci **28**(40): 10075-10080.

Berke, J. D. (2009). "Fast oscillations in cortical-striatal networks switch frequency following rewarding events and stimulant drugs." Eur J Neurosci **30**(5): 848-859.

Berke, J. D., J. T. Breck, et al. (2009). "Striatal versus hippocampal representations during win-stay maze performance." J Neurophysiol **101**(3): 1575-1587.

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

Boyden, E. S., F. Zhang, et al. (2005). "Millisecond-timescale, genetically targeted optical control of neural activity." Nat Neurosci **8**(9): 1263-1268.

Braun, A. A., D. L. Graham, et al. (2012). "Dorsal striatal dopamine depletion impairs both allocentric and egocentric navigation in rats." Neurobiol Learn Mem.

Brown, H. D., J. E. McCutcheon, et al. (2011). "Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum." Eur J Neurosci **34**(12): 1997-2006.

Brown, P. (2007). "Abnormal oscillatory synchronisation in the motor system leads to impaired movement." Curr Opin Neurobiol **17**(6): 656-664.

Buhusi, C. V. and W. H. Meck (2005). "What makes us tick? Functional and neural mechanisms of interval timing." Nat Rev Neurosci **6**(10): 755-765.

Buzsaki, G. (2004). "Large-scale recording of neuronal ensembles." Nat Neurosci **7**(5): 446-451.

Buzsaki, G. and A. Draguhn (2004). "Neuronal oscillations in cortical networks." Science **304**(5679): 1926-1929.

Buzsaki, G. and X. J. Wang (2012). "Mechanisms of gamma oscillations." Annu Rev Neurosci **35**: 203-225.

Calabresi, P. and M. Di Filippo (2008). "ACh/dopamine crosstalk in motor control and reward: a crucial role for alpha 6-containing nicotinic receptors?" Neuron **60**(1): 4-7.
Calabresi, P., R. Maj, et al. (1992). "Long-term synaptic depression in the striatum: physiological and pharmacological characterization." J Neurosci **12**(11): 4224-4233.
Calabresi, P., B. Picconi, et al. (2007). "Dopamine-mediated regulation of corticostriatal synaptic plasticity." Trends Neurosci **30**(5): 211-219.
Calaminus, C. and W. Hauber (2007). "Intact discrimination reversal learning but slowed responding to reward-predictive cues after dopamine D1 and D2 receptor blockade in the nucleus accumbens of rats." Psychopharmacology (Berl) **191**(3): 551-566.
Calaminus, C. and W. Hauber (2009). "Modulation of behavior by expected reward magnitude depends on dopamine in the dorsomedial striatum." Neurotox Res **15**(2): 97-110.
Chuhma, N., K. F. Tanaka, et al. (2011). "Functional connectome of the striatal medium spiny neuron." J Neurosci **31**(4): 1183-1192.
Clark, J. J., S. G. Sandberg, et al. (2010). "Chronic microsensors for longitudinal, subsecond dopamine detection in behaving animals." Nat Methods **7**(2): 126-129.
Cohen, J. Y., S. Haesler, et al. (2012). "Neuron-type-specific signals for reward and punishment in the ventral tegmental area." Nature **482**(7383): 85-88.
Courtemanche, R., N. Fujii, et al. (2003). "Synchronous, focally modulated beta-band oscillations characterize local field potential activity in the striatum of awake behaving monkeys." J Neurosci **23**(37): 11741-11752.
Darvas, M. and R. D. Palmiter (2009). "Restriction of dopamine signaling to the dorsolateral striatum is sufficient for many cognitive behaviors." Proc Natl Acad Sci U S A **106**(34): 14664-14669.
Day, J. J., J. L. Jones, et al. (2010). "Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs." Biol Psychiatry **68**(3): 306-309.
Day, J. J., M. F. Roitman, et al. (2007). "Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens." Nat Neurosci **10**: 1020-1028.
Day, J. J., M. F. Roitman, et al. (2007). "Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens." Nat Neurosci **10**(8): 1020-1028.
DeCoteau, W. E., C. Thorn, et al. (2007). "Learning-related coordination of striatal and hippocampal theta rhythms during acquisition of a procedural maze task." Proc Natl Acad Sci U S A **104**(13): 5644-5649.
Deisseroth, K., G. Feng, et al. (2006). "Next-generation optical technologies for illuminating genetically targeted brain circuits." J Neurosci **26**(41): 10380-10386.
Dejean, C., G. Arbuthnott, et al. (2011). "Power fluctuations in beta and gamma frequencies in rat globus pallidus: association with specific phases of slow oscillations and differential modulation by dopamine D1 and D2 receptors." J Neurosci **31**(16): 6098-6107.
Denk, W., K. L. Briggman, et al. (2012). "Structural neurobiology: missing link to a mechanistic understanding of neural computation." Nat Rev Neurosci **13**(5): 351-358.
Derdikman, D. and E. I. Moser (2010). "A manifold of spatial maps in the brain." Trends Cogn Sci **14**(12): 561-569.
Dombeck, D. A., A. N. Khabbaz, et al. (2007). "Imaging large-scale neural activity with cellular resolution in awake, mobile mice." Neuron **56**(1): 43-57.
Doupe, A. J., D. J. Perkel, et al. (2005). "Birdbrains could teach basal ganglia research a new song." Trends Neurosci **28**(7): 353-363.

Doupe, A. J., M. M. Solis, et al. (2004). "Cellular, circuit, and synaptic mechanisms in song learning." Ann N Y Acad Sci **1016**: 495-523.
Engel, A. K. and P. Fries (2010). "Beta-band oscillations--signalling the status quo?" Curr Opin Neurobiol **20**(2): 156-165.
Engel, A. K., P. Fries, et al. (2001). "Dynamic predictions: oscillations and synchrony in top-down processing." Nat Rev Neurosci **2**(10): 704-716.
Eschenko, O. and S. J. Mizumori (2007). "Memory influences on hippocampal and striatal neural codes: effects of a shift between task rules." Neurobiol Learn Mem **87**(4): 495-509.
Exley, R., J. M. McIntosh, et al. (2012). "Striatal alpha5 nicotinic receptor subunit regulates dopamine transmission in dorsal striatum." J Neurosci **32**(7): 2352-2356.
Fee, M. S. and C. Scharff (2010). "The songbird as a model for the generation and learning of complex sequential behaviors." ILAR J **51**(4): 362-377.
Fiorillo, C. D., P. N. Tobler, et al. (2003). "Discrete coding of reward probability and uncertainty by dopamine neurons." Science **299**(5614): 1898-1902.
Fiorillo, C. D., P. N. Tobler, et al. (2005). "Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors." Behav Brain Funct **1**(1): 7.
Flagel, S. B., J. J. Clark, et al. (2011). "A selective role for dopamine in stimulus-reward learning." Nature **469**(7328): 53-57.
Floresco, S. B., C. L. Todd, et al. (2001). "Glutamatergic afferents from the hippocampus to the nucleus accumbens regulate activity of ventral tegmental area dopamine neurons." J Neurosci **21**(13): 4915-4922.
Floresco, S. B., A. R. West, et al. (2003). "Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission." Nat Neurosci **6**(9): 968-973.
Foster, D. J., R. G. Morris, et al. (2000). "A model of hippocampally dependent navigation, using the temporal difference learning rule." Hippocampus **10**(1): 1-16.
Foster, D. J. and M. A. Wilson (2006). "Reverse replay of behavioural sequences in hippocampal place cells during the awake state." Nature **440**(7084): 680-683.
Frank, M. J., L. C. Seeberger, et al. (2004). "By carrot or by stick: cognitive reinforcement learning in parkinsonism." Science **306**(5703): 1940-1943.
Gage, G. J., C. R. Stoetzner, et al. (2010). "Selective activation of striatal fast-spiking interneurons during choice execution." Neuron **67**(3): 466-479.
Gan, J. O., M. E. Walton, et al. (2010). "Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine." Nat Neurosci **13**(1): 25-27.
Gerfen, C. R. (1985). "The neostriatal mosaic. I. Compartmental organization of projections from the striatum to the substantia nigra in the rat." J Comp Neurol **236**(4): 454-476.
Gerfen, C. R., T. M. Engber, et al. (1990). "D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons." Science **250**(4986): 1429-1432.
Gerfen, C. R. and D. J. Surmeier (2011). "Modulation of striatal projection systems by dopamine." Annu Rev Neurosci **34**: 441-466.
Gold, J. I. and M. N. Shadlen (2007). "The neural basis of decision making." Annu Rev Neurosci **30**: 535-574.
Goldstein, B. L., B. R. Barnett, et al. (2012). "Ventral striatum encodes past and predicted value independent of motor contingencies." J Neurosci **32**(6): 2027-2036.

Grace, A. A. and B. S. Bunney (1984). "The control of firing pattern in nigral dopamine neurons: burst firing." J Neurosci **4**(11): 2877-2890.
Grace, A. A. and B. S. Bunney (1984). "The control of firing pattern in nigral dopamine neurons: single spike firing." J Neurosci **4**(11): 2866-2876.
Grace, A. A., S. B. Floresco, et al. (2007). "Regulation of firing of dopaminergic neurons and control of goal-directed behaviors." Trends Neurosci **30**(5): 220-227.
Graybiel, A. M. (2005). "The basal ganglia: learning new tricks and loving it." Curr Opin Neurobiol **15**(6): 638-644.
Graybiel, A. M. (2008). "Habits, rituals, and the evaluative brain." Annu Rev Neurosci **31**: 359-387.
Graybiel, A. M. and C. W. Ragsdale, Jr. (1978). "Histochemically distinct compartments in the striatum of human, monkeys, and cat demonstrated by acetylthiocholinesterase staining." Proc Natl Acad Sci U S A **75**(11): 5723-5726.
Graybiel, A. M., C. W. Ragsdale, Jr., et al. (1981). "An immunohistochemical study of enkephalins and other neuropeptides in the striatum of the cat with evidence that the opiate peptides are arranged to form mosaic patterns in register with the striosomal compartments visible by acetylcholinesterase staining." Neuroscience **6**(3): 377-397.
Haber, S. N., J. L. Fudge, et al. (2000). "Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum." J Neurosci **20**(6): 2369-2382.
Headley, D. B. and N. M. Weinberger (2011). "Gamma-band activation predicts both associative memory and cortical plasticity." J Neurosci **31**(36): 12748-12758.
Hernandez, P. J., K. Sadeghian, et al. (2002). "Early consolidation of instrumental learning requires protein synthesis in the nucleus accumbens." Nat Neurosci **5**(12): 1327-1331.
Hikosaka, O., H. Nakahara, et al. (1999). "Parallel neural networks for learning sequential procedures." Trends Neurosci **22**(10): 464-471.
Hikosaka, O., M. Sakamoto, et al. (1989). "Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward." J Neurophysiol **61**(4): 814-832.
Hikosaka, O. and R. H. Wurtz (1985). "Modification of saccadic eye movements by GABA-related substances. II. Effects of muscimol in monkey substantia nigra pars reticulata." J Neurophysiol **53**(1): 292-308.
Hok, V., P. P. Lenck-Santini, et al. (2007). "Goal-related activity in hippocampal place cells." J Neurosci **27**(3): 472-482.
Hok, V., E. Save, et al. (2005). "Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex." Proc Natl Acad Sci U S A **102**(12): 4602-4607.
Hollup, S. A., S. Molden, et al. (2001). "Accumulation of hippocampal place fields at the goal location in an annular watermaze task." J Neurosci **21**(5): 1635-1644.
Howe, M. W., H. E. Atallah, et al. (2011). "Habit learning is associated with major shifts in frequencies of oscillatory activity and synchronized spike firing in striatum." Proc Natl Acad Sci U S A **108**(40): 16801-16806.
Humphries, M. D., M. Khamassi, et al. (2012). "Dopaminergic Control of the Exploration-Exploitation Trade-Off via the Basal Ganglia." Front Neurosci **6**: 9.
Humphries, M. D. and T. J. Prescott (2010). "The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward." Prog Neurobiol **90**(4): 385-417.
Jain, V., H. S. Seung, et al. (2010). "Machines that learn to segment images: a crucial technology for connectomics." Curr Opin Neurobiol **20**(5): 653-666.

Jenkinson, N. and P. Brown (2011). "New insights into the relationship between dopamine, beta oscillations and motor function." Trends Neurosci **34**(12): 611-618.
Jin, D. Z., N. Fujii, et al. (2009). "Neural representation of time in cortico-basal ganglia circuits." Proc Natl Acad Sci U S A **106**(45): 19156-19161.
Joel, D. and I. Weiner (2000). "The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum." Neuroscience **96**(3): 451-474.
Kalenscher, T., C. S. Lansink, et al. (2010). "Reward-associated gamma oscillations in ventral striatum are regionally differentiated and modulate local firing activity." J Neurophysiol **103**(3): 1658-1672.
Kalivas, P. W., L. Churchill, et al. (1993). "GABA and enkephalin projection from the nucleus accumbens and ventral pallidum to the ventral tegmental area." Neuroscience **57**(4): 1047-1060.
Keithley, R. B., M. L. Heien, et al. (2009). "Multivariate concentration determination using principal component regression with residual analysis." Trends Analyt Chem **28**(9): 1127-1136.
Kimchi, E. Y. and M. Laubach (2009). "Dynamic encoding of action selection by the medial striatum." J Neurosci **29**(10): 3148-3159.
Kimchi, E. Y., M. M. Torregrossa, et al. (2009). "Neuronal correlates of instrumental learning in the dorsal striatum." J Neurophysiol **102**(1): 475-489.
Kish, L. J., M. R. Palmer, et al. (1999). "Multiple single-unit recordings in the striatum of freely moving animals: effects of apomorphine and D-amphetamine in normal and unilateral 6-hydroxydopamine-lesioned rats." Brain Res **833**(1): 58-70.
Kita, T., H. Kita, et al. (1985). "Local stimulation induced GABAergic response in rat striatal slice preparations: intracellular recordings on QX-314 injected neurons." Brain Res **360**(1-2): 304-310.
Kravitz, A. V., B. S. Freeze, et al. (2010). "Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry." Nature **466**(7306): 622-626.
Krebs, M. O., F. Trovero, et al. (1991). "Distinct presynaptic regulation of dopamine release through NMDA receptors in striosome- and matrix-enriched areas of the rat striatum." J Neurosci **11**(5): 1256-1262.
Kubota, Y., J. Liu, et al. (2009). "Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum." J Neurophysiol **102**(4): 2142-2160.
Lansink, C. S., P. M. Goltstein, et al. (2009). "Hippocampus leads ventral striatum in replay of place-reward information." PLoS Biol **7**(8): e1000173.
Lansink, C. S., J. C. Jackson, et al. (2012). "Reward Cues in Space: Commonalities and Differences in Neural Coding by Hippocampal and Ventral Striatal Ensembles." J Neurosci **32**(36): 12444-12459.
Lavoie, A. M. and S. J. Mizumori (1994). "Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of rats." Brain Res **638**(1-2): 157-168.
Le Moine, C. and B. Bloch (1995). "D1 and D2 dopamine receptor gene expression in the rat striatum: sensitive cRNA probes demonstrate prominent segregation of D1 and D2 mRNAs in distinct neuronal populations of the dorsal and ventral striatum." J Comp Neurol **355**(3): 418-426.
Leventhal, D. K., G. J. Gage, et al. (2012). "Basal ganglia beta oscillations accompany cue utilization." Neuron **73**(3): 523-536.

Lichtman, J. W. and J. R. Sanes (2008). "Ome sweet ome: what can the genome tell us about the connectome?" Curr Opin Neurobiol **18**(3): 346-353.
Lisman, J. E. and A. A. Grace (2005). "The hippocampal-VTA loop: controlling the entry of information into long-term memory." Neuron **46**(5): 703-713.
Louie, K. and M. A. Wilson (2001). "Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep." Neuron **29**(1): 145-156.
Luo, A. H., P. Tahsili-Fahadan, et al. (2011). "Linking context with reward: a functional circuit from hippocampal CA3 to ventral tegmental area." Science **333**(6040): 353-357.
Margolis, E. B., H. Lock, et al. (2006). "The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons?" J Physiol **577**(Pt 3): 907-924.
Markram, H. (2012). "The human brain project." Sci Am **306**(6): 50-55.
Masimore, B., N. C. Schmitzer-Torbert, et al. (2005). "Transient striatal gamma local field potentials signal movement initiation in rats." Neuroreport **16**(18): 2021-2024.
Matsuda, W., T. Furuta, et al. (2009). "Single nigrostriatal dopaminergic neurons form widely spread and highly dense axonal arborizations in the neostriatum." J Neurosci **29**(2): 444-453.
Matsumoto, M. and O. Hikosaka (2009). "Two types of dopamine neuron distinctly convey positive and negative motivational signals." Nature **459**(7248): 837-841.
Miller, E. K. and J. D. Cohen (2001). "An integrative theory of prefrontal cortex function." Annu Rev Neurosci **24**: 167-202.
Mink, J. W. (1996). "The basal ganglia: focused selection and inhibition of competing motor programs." Prog Neurobiol **50**(4): 381-425.
Miyachi, S., O. Hikosaka, et al. (2002). "Differential activation of monkey striatal neurons in the early and late stages of procedural learning." Exp Brain Res **146**(1): 122-126.
Miyachi, S., O. Hikosaka, et al. (1997). "Differential roles of monkey striatum in learning of sequential hand movement." Exp Brain Res **115**(1): 1-5.
Mogenson, G. J., D. L. Jones, et al. (1980). "From motivation to action: functional interface between the limbic system and the motor system." Prog Neurobiol **14**(2-3): 69-97.
Morris, G., D. Arkadir, et al. (2004). "Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons." Neuron **43**(1): 133-143.
Morris, G., A. Nevet, et al. (2006). "Midbrain dopamine neurons encode decisions for future action." Nat Neurosci **9**(8): 1057-1063.
Nair-Roberts, R. G., S. D. Chatelain-Badie, et al. (2008). "Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat." Neuroscience **152**(4): 1024-1031.
Nicola, S. M., I. A. Yun, et al. (2004). "Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task." J Neurophysiol **91**(4): 1840-1865.
Nordeen, K. W. and E. J. Nordeen (1997). "Anatomical and synaptic substrates for avian song learning." J Neurobiol **33**(5): 532-548.
Obeso, J. A., C. Marin, et al. (2008). "The basal ganglia in Parkinson's disease: current concepts and unexplained observations." Ann Neurol **64 Suppl 2**: S30-46.
Overton, P. and D. Clark (1992). "Iontophoretically administered drugs acting at the N-methyl-D-aspartate receptor modulate burst firing in A9 dopamine neurons in the rat." Synapse **10**(2): 131-140.

Parthasarathy, H. B. and A. M. Graybiel (1997). "Cortically driven immediate-early gene expression reflects modular influence of sensorimotor cortex on identified striatal neurons in the squirrel monkey." J Neurosci **17**(7): 2477-2491.
Popescu, A. T., D. Popa, et al. (2009). "Coherent gamma oscillations couple the amygdala and striatum during learning." Nat Neurosci **12**(6): 801-807.
Redish, A. D., S. Jensen, et al. (2008). "A unified framework for addiction: vulnerabilities in the decision process." Behav Brain Sci **31**(4): 415-437; discussion 437-487.
Reid, R. C. (2012). "From functional architecture to functional connectomics." Neuron **75**(2): 209-217.
Reynolds, J. N. and J. R. Wickens (2002). "Dopamine-dependent plasticity of corticostriatal synapses." Neural Netw **15**(4-6): 507-521.
Richfield, E. K., J. B. Penney, et al. (1989). "Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system." Neuroscience **30**(3): 767-777.
Robinson, D. L., B. J. Venton, et al. (2003). "Detecting subsecond dopamine release with fast-scan cyclic voltammetry in vivo." Clin Chem **49**(10): 1763-1773.
Salamone, J. D. and M. Correa (2002). "Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine." Behav Brain Res **137**(1-2): 3-25.
Salamone, J. D., M. Correa, et al. (2007). "Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits." Psychopharmacology (Berl) **191**(3): 461-482.
Sanes, J. N. and J. P. Donoghue (1993). "Oscillations in local field potentials of the primate motor cortex during voluntary movement." Proc Natl Acad Sci U S A **90**(10): 4470-4474.
Sargolini, F., M. Fyhn, et al. (2006). "Conjunctive representation of position, direction, and velocity in entorhinal cortex." Science **312**(5774): 758-762.
Schultz, W. (2002). "Getting formal with dopamine and reward." Neuron **36**(2): 241-263.
Schultz, W. (2011). "Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs." Neuron **69**(4): 603-617.
Schultz, W., P. Apicella, et al. (1993). "Reward-related activity in the monkey striatum and substantia nigra." Prog Brain Res **99**: 227-235.
Schultz, W., P. Apicella, et al. (1992). "Neuronal activity in monkey ventral striatum related to the expectation of reward." J Neurosci **12**(12): 4595-4610.
Schultz, W., P. Dayan, et al. (1997). "A neural substrate of prediction and reward." Science **275**(5306): 1593-1599.
Setlow, B., G. Schoenbaum, et al. (2003). "Neural encoding in ventral striatum during olfactory discrimination learning." Neuron **38**(4): 625-636.
Sharp, P. E. and C. Green (1994). "Spatial correlates of firing patterns of single cells in the subiculum of the freely moving rat." J Neurosci **14**(4): 2339-2356.
Shen, W., M. Flajolet, et al. (2008). "Dichotomous dopaminergic control of striatal synaptic plasticity." Science **321**(5890): 848-851.
Shepard, P. D. and B. S. Bunney (1991). "Repetitive firing properties of putative dopamine-containing neurons in vitro: regulation by an apamin-sensitive Ca(2+)-activated K+ conductance." Exp Brain Res **86**(1): 141-150.
Shepherd, G., Ed. (2004). The Synaptic Organization of the Brain. New York, Oxford University Press.

Stalnaker, T. A., G. G. Calhoon, et al. (2010). "Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum." Front Integr Neurosci **4**: 12.

Stalnaker, T. A., G. G. Calhoon, et al. (2012). "Reward prediction error signaling in posterior dorsomedial striatum is action specific." J Neurosci **32**(30): 10296-10305.

Suri, R. E. and W. Schultz (1998). "Learning of sequential movements by neural network model with dopamine-like reinforcement signal." Exp Brain Res **121**(3): 350-354.

Surmeier, D. J., J. Ding, et al. (2007). "D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons." Trends Neurosci **30**(5): 228-235.

Sutton, R. S., & Barto, A.G., Ed. (1998). Reinforcement Learning: An Introduction. Cambridge, MA, MIT Press.

Thorn, C. A., H. Atallah, et al. (2010). "Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning." Neuron **66**(5): 781-795.

Tobler, P. N., C. D. Fiorillo, et al. (2005). "Adaptive coding of reward value by dopamine neurons." Science **307**(5715): 1642-1645.

Tort, A. B., M. A. Kramer, et al. (2008). "Dynamic cross-frequency couplings of local field potential oscillations in rat striatum and hippocampus during performance of a T-maze task." Proc Natl Acad Sci U S A **105**(51): 20517-20522.

van der Meer, M. A., T. Kalenscher, et al. (2010). "Integrating early results on ventral striatal gamma oscillations in the rat." Front Neurosci **4**: 300.

van der Meer, M. A. and A. D. Redish (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." Front Integr Neurosci **3**: 9.

van der Meer, M. A. and A. D. Redish (2011). "Theta phase precession in rat ventral striatum links place and reward information." J Neurosci **31**(8): 2843-2854.

Voorn, P., L. J. Vanderschuren, et al. (2004). "Putting a spin on the dorsal-ventral divide of the striatum." Trends Neurosci **27**(8): 468-474.

Waelti, P., A. Dickinson, et al. (2001). "Dopamine responses comply with basic assumptions of formal learning theory." Nature **412**(6842): 43-48.

Walsh, M. M. and J. R. Anderson (2011). "Learning from delayed feedback: neural responses in temporal credit assignment." Cogn Affect Behav Neurosci **11**(2): 131-143.

Wang, L. P., F. Li, et al. (2011). "NMDA receptors in dopaminergic neurons are crucial for habit learning." Neuron **72**(6): 1055-1066.

Wang, X. J. (2010). "Neurophysiological and computational principles of cortical rhythms in cognition." Physiol Rev **90**(3): 1195-1268.

Wespatat, V., F. Tennigkeit, et al. (2004). "Phase sensitivity of synaptic modifications in oscillating cells of rat visual cortex." J Neurosci **24**(41): 9067-9075.

Whishaw, I. Q. and S. B. Dunnett (1985). "Dopamine depletion, stimulation or blockade in the rat disrupts spatial navigation and locomotion dependent upon beacon or distal cues." Behav Brain Res **18**(1): 11-29.

Wickens, J. R., J. C. Horvitz, et al. (2007). "Dopaminergic mechanisms in actions and habits." J Neurosci **27**(31): 8181-8183.

Wightman, R. M., M. L. Heien, et al. (2007). "Dopamine release is heterogeneous within microenvironments of the rat nucleus accumbens." Eur J Neurosci **26**(7): 2046-2054.

Wilson, M. A. and B. L. McNaughton (1993). "Dynamics of the hippocampal ensemble code for space." Science **261**(5124): 1055-1058.

Witten, I. B., E. E. Steinberg, et al. (2011). "Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement." Neuron **72**(5): 721-733.

Wu, Q., M. E. Reith, et al. (2001). "Preferential increases in nucleus accumbens dopamine after systemic cocaine administration are caused by unique characteristics of dopamine neurotransmission." J Neurosci **21**(16): 6338-6347.

Yin, H. H. and B. J. Knowlton (2006). "The role of the basal ganglia in habit formation." Nat Rev Neurosci **7**(6): 464-476.

Yin, H. H., B. J. Knowlton, et al. (2004). "Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning." Eur J Neurosci **19**(1): 181-189.

Yin, H. H., B. J. Knowlton, et al. (2005). "Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning." Eur J Neurosci **22**(2): 505-512.

Yin, H. H., B. J. Knowlton, et al. (2006). "Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning." Behav Brain Res **166**(2): 189-196.

Yin, H. H., S. B. Ostlund, et al. (2005). "The role of the dorsomedial striatum in instrumental conditioning." Eur J Neurosci **22**(2): 513-523.

Yun, I. A., K. T. Wakabayashi, et al. (2004). "The ventral tegmental area is required for the behavioral and nucleus accumbens neuronal firing responses to incentive cues." J Neurosci **24**(12): 2923-2933.

Zhang, H. and D. Sulzer (2003). "Glutamate spillover in the striatum depresses dopaminergic transmission by activating group I metabotropic glutamate receptors." J Neurosci **23**(33): 10585-10592.

Zhang, L., W. M. Doyon, et al. (2009). "Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum." Mol Pharmacol **76**(2): 396-404.

Zhou, Q. Y. and R. D. Palmiter (1995). "Dopamine-deficient mice are severely hypoactive, adipsic, and aphagic." Cell **83**(7): 1197-1209.

Zweifel, L. S., J. G. Parker, et al. (2009). "Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior." Proc Natl Acad Sci U S A **106**(18): 7281-7288.