Neural coding within striatal subregions: a shared framework with specializations

By

Alaina Catherine Case

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Neuroscience)
in the University of Michigan
2014**

Doctoral Committee:

Associate Professor Joshua D. Berke, Chair
Professor J. Wayne Aldridge
Professor Kent C. Berridge
Associate Professor Michael M. A. Sutton

Dedication:

To my brother, Cory: I just want to understand your brain

Table of Contents

List of Figures

Legend:
A: Previous Outcome Only cells. From a new 3-factor regression examining Previous Outcome, Goal and Cue, these are cells that reached significance for the factor Previous Outcome Only, plotted 6 seconds around each of the behavioral events. The filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps. All 5 brain regions significantly encode whether or not the previous trial was rewarded, prior to the tone playing in the current trial.
B: Previous Outcome Firing Rate Index. From the cells in A, the trial type that had the highest firing rate is plotted as an index across time. Trials are identified by whether or not the previous outcome was correct (Previous Correct - Previous Incorrect). The filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Inset shows the area under the curve of the line plotted in B, from -3 to 1s before Tone.
C: Previous Outcome Firing Rate Index. From the cells in A, the trial type that had the highest firing rate is plotted as an index across time. Trials are identified by whether or not the previous outcome was correct (Previous Correct - Previous Incorrect). The filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Inset shows the area under the curve of the line plotted in B, from -3 to 1s before reward port in.
D: Previous Outcome plotted on Sated and Restricted days only. Same analysis used in 14A.

Legend:
A: Encoding by Region. The 3-Factor regression Action-Goal-Cue is plotted for all of the relevant factors for the cells in just the DLS, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
B: Individual Cell Plots. Each cell that was recorded from in the DLS is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Goal' encoding at reward port in, and this sorting order remained consistent across all behavioral events.
C: Event Overlap. The Venn Diagram represents Action-Goal-Outcome encoding during the events Nose Side In (-0.5 to 0.5 seconds for action), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Legend:
A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the Core, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
B: Individual Cell Plots. Each cell that was recorded from in the Core is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.
C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Legend:
A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the Shell, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
B: Individual Cell Plots. Each cell that was recorded from in the Shell is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.
C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Legend:
A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the OFC, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached

significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
B: Individual Cell Plots. Each cell that was recorded from in the OFC is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.
C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Abbreviations:

ac, anterior commissure;

ACd, dorsal anterior cingulate cortex;

AId, dorsal agranular insular cortex;

AIv, ventral agranular insular cortex;

CeM, central medial thalamic nucleus;

CL, central lateral thalamic nucleus;

DA, dopamine;

D1, dopamine D1 containing receptor;

D2, dopamine D2 containing receptor;

DS dorsal striatum;

DLS, Dorsal Lateral Striatum;

DMS, Dorsal Medial Striatum;

FSI, fast spiking interneuron;

GPe, external globus pallidus;

GPi, internal globus pallidus;

IL, infralimbic cortex;

IMD, intermediodorsal thalamic nucleus;

MD, mediodorsal thalamic nucleus;

MSN, medium spiny neuron;

NAc, Nucleus Accumbens;

OFC, Orbitofrontal Cortex;

PC, paracentral thalamic nucleus;

PFC, prefrontal cortex;

PLd, dorsal prelimbic cortex;

PLv, ventral prelimbic cortex;

PV, paraventricular thalamic nucleus;

RL, Reinforcement-Learning;

RPE, reward-prediction error;

SMC, sensorimotor cortex.

SNc, substantia nigra pars compacta;

SNr, substantia nigra pars reticulata;

STN, subthalamic nucleus;

TD, temporal difference;

Abstract:

Parallel cortical-striatal loop circuits participate in distinct forms of decision-making that make use of different representations. To investigate how these circuits process information that translates internal state into action selection, we designed a behavioral task that uses hunger and thirst to manipulate motivational states. Rats had to follow an instructed cue guiding them to either food or water reward on a majority of trials, and were only rewarded if they chose correctly. On 25% of trials the rats were allowed to freely choose the reward in order to test how internal state affects choice behavior. Behavioral performance on the task demonstrates that internal state accurately guides decisions on both free choice and instructed choice trials.

To investigate how the cortical-striatal loop circuits process information in this unique behavioral task, we compared single-neuron activity across multiple striatal subregions, together with orbitofrontal cortex (OFC), to tease apart coding of overt movements, expected goals, and internal motivational states. Recordings were made from 7 rats, in 2,345 cells. Overall, neural activity evolved in a remarkably similar way across all recorded areas, with each region showing significant encoding of movement direction, expected goal, reward outcome and reward history.

Within each striatal subregion we also found neurons encoding a unique combination of task elements. Neurons responsive to the combination of movement direction and current goal were found selectively in dorsomedial striatum (DMS), thought to be important for action-outcome associations. Tone response was found exclusively in the dorsolateral striatum (DLS) and nucleus accumbens shell. All recorded areas responded in anticipation of the upcoming goal choice, while no regions directly represented the correct cue, prior to receiving reward. Only the dorsomedial striatum showed changes in firing rate due to internal motivational state, while all other brain regions specialized in reward-identity coding. Additionally, although each region significantly encoded Action, Goal, Outcome and Reward History, the relative contribution made by each region varied, demonstrating each regions unique specialization.

These results indicate that parallel cortical-striatal circuits share a common information processing framework, with specific neuronal subpopulations supporting distinct specializations.

Chapter 1: Internal State and the Cortico-Basal Ganglia Circuitry

When hungry, how does the brain know to go to the cupboard and get a bag of Oreos, rather than head to the fridge and pour a glass of milk? Internal motivational states dictate what appropriate actions to take when standing in the kitchen, deciding between the fridge and the cupboard. How the neural correlates of translating hunger and thirst into action-selection are still unclear.

Flexible decision-making uses both external sensory information and internal motivational drives (or states) to influence the choices that an animal makes. External cues may guide reward selection, but internal properties, such as hunger and thirst, are salient motivational factors in a natural environment. How does the brain encode these internal states that then drive motivated actions? Neuroscientists have been studying rewarded behavior for decades, but most studies rely on externally driven motivational manipulations, such as reward probability or reward magnitude. The cortico-basal ganglia circuitry plays a crucial role in integrating information about reward predictions with planning external movements used to obtain rewards. This study uses behavior and electrophysiology to examine and compare the neural encoding of internal states in this circuitry, specifically, the orbitofrontal cortex (OFC), the dorsal striatum (DS) and the ventral striatum (VS).

Dissociating the role of external vs. internal information lends insight into various diseases, such as eating disorders, depression, OCD and obesity. The same circuit

implicated in reward and actions are implicated in these disorders. A common feature in all is that some feature of internal state is misrepresented, causing actions that are inflexible and harmful. The underlying brain structures involved show that irregular brain activity may underlie these disorders. OCD patients show abnormal metabolic activity in the orbitofrontal cortex, the anterior cingulate/caudal medial prefrontal cortex, and the caudate nucleus (Graybiel, 2000). Studies done examining anorexia nervosa have found abnormalities in distinct behavioral tasks and neural circuitry involved in motivation (Zastrow et al., 2009). Human literature also suggests that the underlying pathophysiology of both Depression and OCD center on the prefrontal-basal ganglia system, including the orbitofrontal cortex and both dorsal and ventral striatum (Haber and Brucker, 2009; Drevets, 2000).

Studies examining obesity in humans found that the consumption of palatable foods involves an increase in activation in the right lateral OFC, frontal operculum and insula, indicating that in humans, the OFC is responsible for encoding subjective reward value related to food consumption (O'Doherty et al., 2002). Studies that took into account intrinsic physiological states found that, in contrast to controls, obese women had selective activation in the dorsal striatum, lateral orbitofrontal cortex as well as other frontal cortical regions (Rothemund et al., 2007). It is clear that the basal ganglia and orbitofrontal cortex are involved in internal motivation, decision-making, actions, predicting reward value, and are sensitive to food-related cues. Patients that show a deficit in monitoring and controlling their internal state show a concurrent change in activation in the cortico-basal ganglia circuitry.

The field of decision-making is a wide field of research, and theories explaining

animal behavior come from areas such as behavioral economics, machine learning and classical psychology. Although the behavioral task used in this study does not fall neatly into any one specific category, it is important to discuss the relevant research related to instrumental learning such as goal-directed behavior and stimulus-response behavior, as well as the role of classical conditioning through Pavlovian association, since all of these aspects of behavior are crucial to any decision-making discussion.

Pavlovian Conditioning

One form of learning and decision-making is Classical Conditioning. In this form of learning, an animal learns the relationship between a stimulus and the outcome, so that the paired stimuli can then come to release an action (Niv et al., 2006). Unconditioned responses (UR) are evolutionary responses such as salivating, freezing, approach or fleeing. An animal can learn to associate previously neutral stimuli (such as a tone or a light) with a rewarding event (such as food). The neutral stimulus is initially not capable of eliciting a UR- there is no natural response for a rat when it hears a non-threatening tone, or sees a flashing light. Through repeated pairings of the neutral stimulus with a reward, the previously neutral stimulus then becomes a conditioned stimulus (CS). It does so by evoking the conditioned response (CR) that was appropriate to the reward. The important aspect of Pavlovian conditioning is that the animal can learn this association without his behavior actually being the cause of receiving the reward. The conditioned response develops because of an association that forms between a representation of the CS and the US. The connections between the ventral striatum and amygdala are critical for Pavlovian Behavior (Cardinal et al., 2002a).

In rodents, one way to measure Pavlovian associations is through Pavlovian conditioned approach, which is measured by approach to a light or lever that predicts a reward (sign tracking), or approach to a goal port where food is delivered (goal tracking) (Flagel et al., 2011). Another way Pavlovian associations can be indirectly measured is by their effect on instrumental behavior. In this procedure, a rodent learns the Pavlovian association between a light and the delivery of a sucrose pellet. It is then presented with a lever that, after pressing it, delivers a sucrose reward (an instrumental response). When presented with the light at the same time as the lever, the rat will enhance responding on that lever, due to the motivating effect of Pavlovian instrumental transfer (Lovibond, 1983; Estes, 1948; Cardinal et al., 2002a). The stimulus learned through Pavlovian associations adds to the instrumental effect of lever responding.

Instrumental Learning

Instrumental learning involves forming a relationship between an animal's action and the reinforcing outcome. Repeated delivery of a reward serves to strengthen the association between an environmental stimuli and a particular response. The key difference between the Pavlovian system and instrumental system is that in instrumental conditioning, the agent learns to select specific actions that will increase the probability of reward (Skinner 1938; Thorndike 1911), while Pavlovian conditioning does not require a direct action to receive reward. An instrumental response can be any type of movement or action, and is not necessarily an evolutionarily specific behavior related to the reward association. On the other hand, Pavlovian associations evoke reward-specific responses that are evolutionarily specific.

A series of studies utilizing instrumental learning procedures with various tests on the nature of goal representation have determined that instrumental decision-making is driven by either a goal-directed or a stimulus-response mechanisms (Balleine and Dickinson, 1998; Balleine et al., 2009). The main procedure used to test how a behavior is being represented is called 'Outcome Devaluation'. In the first phase of training, an animal learns that a certain action leads to a reward. The value of the reward is then reduced, either by pairing the reward with feeling sick, or feeding the reward to satiety. The animal is then tested in the previous behavioral paradigm without giving any reward. If the animal reduces responding to the previously rewarded action, then the behavior is thought of as goal directed. If the animal continues to respond in the same manner before devaluing the reward, then the behavior is thought to be habitual, because then his actions are not being driven by the specific goal, since that goal is no longer desired or consumed.

In order to understand how these different aspects of behavior and decision-making are represented in the brain, and translated into action selection, it is important to understand the neurophysiology of one circuit involved in reward and action selection: the Basal Ganglia.

Basal Ganglia Circuitry

The basal ganglia work together with the cortex to orchestrate and execute planned movements by integrating aspects of goal-directed behavior that include elements of motivation, emotion and cognition (Mink, 1996; Balleine et al., 2007). The

general term ‘basal ganglia’ (BG) includes the striatum, the globus pallidus (external-GPe and internal -GPi) the ventral pallidum (VP) the substantia nigra (pars compacta, SNc and pars reticulata, SNr) the ventral tegmental area (VTA) and the subthalmic nucleus (STN). The striatum is the largest part of the basal ganglia, and is the primary input nucleus, receiving excitatory inputs from the thalamus and cortex (Figure 1A). The STN is the other major input structure, and receives connections from the cortex and thalamus. The major output nuclei are the SNr and GPi, which receive input from the striatum, STN and GP, and project to the thalamus and back to the cortex as well as other brainstem nuclei (Figure 1A) (Mink, 1996; Haber, 2003; Grillner et al., 2005).

The striatum can be broadly divided into 3 main components: the dorsolateral (DLS), dorsomedial (DMS) and ventral striatum (VS) (Figure 2B) (Joel and Weiner, 2000; Yin and Knowlton, 2006). Divisions in the dorsal striatum are delineated mainly by the extent of cortical inputs, with the DLS receiving primary motor and somatosensory cortical inputs, and the DMS receiving inputs from the association cortices (Stanton et al., 1988).

The ventral striatum includes the nucleus accumbens (NAc), and generally constitutes the remaining area of the striatum. The two subregions of the NAc are the Core and Shell (Zaborszky et al., 1985; Heimer et al., 1991; Groenewegen et al., 1999). The Core is a dense region of cells surrounding the anterior commissure. Staining for acetylcholinesterase (AChE) and calbindin show clear delineation between the core and shell, with light staining in the core for AChE and dark staining in the shell, while the opposite pattern is true for calbindin staining (Zaborszky et al., 1985; Zahm and Brog, 1992; Jongen-Relo et al., 1994; Groenewegen et al., 1999).

The majority of the striatum is comprised of medium spiny neurons (MSNs) and are the projections neurons of the striatum (Tepper and Bolam, 2004; Matamales et al., 2009). These are GABA-containing neurons that are often quiet, due to their intrinsic membrane properties (Wilson and Bowman, 2004). Activation often requires strong input from the cortex and/or thalamus. When activated, their activity acts to reduce the tonically active downstream targets, and through specific activation of different pathways, the striatum allows action selection through disinhibition (Deniau and Chevalier, 1985; Mink, 1996).

Other striatal neurons include GABAergic interneurons that co-express parvalbumin and are called 'fast-spiking interneurons' (FSIs) (Kawaguchi, 1993), (Kawaguchi, 1995). These fast spiking interneurons act to inhibit local microcircuits and provide more precise control over action selection, by suppressing unwanted actions (Kita et al., 1990), (Parthasarathy and Graybiel, 1997), (Gage et al., 2010). Interneurons compromise around 5% of the striatal neurons, but have broad reaching effects. FSIs form an inhibitory microcircuit in the striatum (Parthasarathy and Graybiel, 1997), (Mallet et al., 2005). FSIs receive glutamatergic input from cortical pyramidal neurons, and in some estimations, it is thought that a single FSI inhibits between 135-541 MSNs (Koós and Tepper, 1999), in both direct and indirect pathways (Kita, 1993; Bennett and Bolam, 1994; ;Gittis et al., 2010; Gittis et al., 2011; Planert et al., 2010). Additional local inhibitory circuits arise from lateral projections of MSNs onto other MSNs, and although this connectivity may be weak and sparse (Kawaguchi et al., 1989; Koos et al., 2004; Jaeger et al., 1994; Taverna et al., 2008) the sum total inhibitory network could form strong inhibition (Chuhma et al., 2011).

Direct vs. Indirect Pathways:

There are two main pathways, through which information flows through the basal ganglia, which have opposite effects on their target neurons (Figure 1B) (Albin et al., 1989; DeLong, 1990). The 'direct' pathway has a disinhibitory effect on targets, acting to promote motor output, while the 'indirect' pathway increases inhibition on target neurons, suppressing action (Chevalier and Deniau, 1990; Mink, 1996). The direct pathway is thought to contain mostly D1 type receptor neurons that co-localize with the neuropeptides substance P and dynorphin. These receptors subtypes are preferentially located on MSNs that project to the SNr and GPi through monosynaptic connections (the so called striatonigral pathway). The indirect pathway is thought to contain mostly D2 receptor neurons and contain enkephalin as well as adenosine A2A receptors. The indirect pathway has a multisynaptic route that synapses on the GPe and STN before reaching the SNr and GPi (the so called striatopallidal pathway) (Figure 1C) (Surmeier et al., 1996).

Using EGFP to tag either D1-like or D2-like receptor promoters on transgenic mice, Bertran-Gonzalez (Bertran-Gonzalez et al., 2008) was able to show that only 5% of the MSNs in the dorsal striatum express both dopamine receptor subtypes, demonstrating a distinct segregation of these two pathways. Additionally, new techniques in tract tracing show conclusively that sensory cortical and limbic structures preferentially innervate the direct pathway, while motor cortex preferentially target the indirect pathway (Wall et al., 2013). The thalamostriatal input, dopamine input and other specific

cortical layer input targets both pathways, while the amygdala shows stronger innervation onto the direct pathway.

The classic model of basal ganglia function proposed that the direct and indirect pathways opposed each other, but more recent studies demonstrate activation of both pathways during a contraversive movement (Cui et al., 2013). Additional proof of principle between activation of the direct and indirect pathways in initiating and inhibiting movement was most recently shown using optogenetics to evoke or inhibit a locomotor response in mice (Kravitz et al., 2010; Kravitz et al., 2012; Tai et al., 2012). Other theories for the direct and indirect pathway propose separate roles in reward related action selection. The direct pathway may facilitate previously rewarded actions, while the indirect pathway may act to suppress previously unrewarded actions (Bromberg-Martin et al., 2010; Frank et al., 2004; Kravitz et al., 2012).

Additional distinctions between the direct and indirect pathway can be seen in their differential response to dopamine, due to differences in dopamine receptor affinities. Different levels of dopamine affect the relative excitability of the two different pathways. The D1-direct pathway is more excitable with a high dopamine level concentration, leading to increased D1-excitation. The D2-indirect pathway is more excitable during low levels of dopamine concentration, due to the fact that the normal suppression of neuronal excitability is now reduced (Albin et al., 1989; Surmeier et al., 2010).

Dorsal Striatum

The Striatum, in general, is the largest component of the basal ganglia circuitry, and in addition to the STN, is the only input nuclei in the basal ganglia. The Dorsal

Striatum receives excitatory glutamatergic inputs from almost the entire cerebral cortex, including motor, sensory, association and limbic areas. Additional inputs come from the limbic system and thalamus, as well as dopaminergic inputs from the midbrain (Sesack et al., 2003). The thalamic inputs come from the midline and intralaminar nuclei. The main outputs of the basal ganglia are the substantia nigra and the globus pallidus, which then project to the thalamus (and then on to the cortex) and to pre-motor areas of the brain stem (Groenewegen, 2003; Redgrave et al., 1999). The predominant class of neuron is the medium spiny neuron (MSN), which receives and integrates most of the inputs from the cerebral cortex and thalamus. These MSNs contain the inhibitory neurotransmitter GABA and project to the pallidum and substantia nigra (Gerfen and Wilson, 1996).

The dorsal striatum can be divided into two subregions based on functional connectivity and behavioral studies (Figure 2A). The dorsal medial striatum (DMS) receives input from the associative areas of the prefrontal cortex while the dorsal lateral striatum (DLS) receives inputs from the sensorimotor cortex (Alexander et al., 1986), (Groenewegen et al., 1990). Balleine and colleagues have demonstrated from a series of studies that the dorsal medial striatum (DMS) is involved in goal-directed behavior, while the dorsal lateral striatum (DLS) is important for habitual behavior (Yin et al., 2008a; Balleine et al., 2009). The specific behavioral roles of DMS and DLS will be discussed in upcoming sections.

Differences between the DMS and DLS can also be seen through their differential expression of specific types of learning. Within the striatum, D1 and D2 pathways differ in their expression of synaptic plasticity through long term depression (LTD) or long term potentiation (LTP). LTP can be induced by the activation of D1-like dopamine

receptors and NMDA glutamate receptors (Partridge et al., 2000; Kerr and Wickens, 2001). Blockade of NMDA receptors in the DMS prevents learning action-outcome contingencies (Yin et al., 2005b). LTP is usually found in the DMS. In contrast, the DLS usually displays LTD, which requires the activation of D2-like dopamine receptors, group I metabotropic glutamate receptors and L-type calcium channels (Gerderman 2002). The separation of D1 and D2 containing neurons, paired with their differential downstream projections provide the neural framework for learning and reward-guided action-selection.

Nucleus Accumbens

Inputs: The dorsal striatum and both regions of the ventral striatum share common inputs from the neocortex, thalamus and dopaminergic cells in the brainstem. The nucleus accumbens receives inputs from the hippocampus and amygdala (Kelley and Domesick, 1982; Groenewegen et al., 1987). The entorhinal cortex, both medial and lateral, send projections to the medial and lateral divisions of both core and shell (Totterdell and Meredith, 1997). Within the prefrontal cortex, the strongest inputs to core and shell (as well as to the dorsal striatum) are from the medial prefrontal cortex, which include the prelimbic area, medial orbital area and infralimbic area (Ding et al., 2001). The infralimbic area only projects to the medial shell. Projections from the dorsal and ventral agranular insula project to the caudo-lateral and rostro-lateral core and shell (Hoover and Vertes, 2011).

Core Connectivity

The connectivity and morphology of the Core is very similar to the dorsomedial striatum (Humphries and Prescott, 2010). The Core contains two populations of MSNs that express D1 or D2 receptors (Lu et al., 1998), whose main targets are restricted to nuclei within the basal ganglia. The overlap between D1 and D2 containing MSNs within the same neuron remains low, like the dorsal striatum, at 6% (Bertran-Gonzalez et al., 2008). Projections from the core target subdivisions of the SNr (Deniau et al., 1996; Deniau et al., 1994), and are primarily formed by D1-dominant MSNs (Zhou et al., 2003; Lu et al., 1998). Additionally, the core projects to the dorsolateral ventral pallidum (Heimer et al., 1991; Maurice et al., 1997), which then projects to the medial STN then on to the dorsomedial SNr, which is the target of the direct projections from the core. The projections from core to the VP consist of all the MSNs with D2 receptors (Lu et al., 1998; Zhou et al., 2003), however there is also a sub-population of D1 MSNs that also project to the VP. This dorsolateral VP projects back to the core (Hakan et al., 1992; Groenewegen et al., 1993).

Shell Connectivity

The shell distinguishes itself by having projections to structures outside of the basal ganglia. Within the shell, the dorsomedial area projects directly to regions of the lateral hypothalamus (LH) and lateral pre-optic area. The ventromedial shell projects to the adjacent areas of those same structures, as well as to the parabrachial nucleus, the periacqueductal grey and adjacent areas (Mogenson et al., 1983; Zahm and Brog, 1992), (Usuda et al., 1998). The lateral shell projections stay within the basal ganglia (Usuda et al., 1998). Additionally, the Shell has a higher proportion of MSNs that express both D1-

like and D2-like receptors (17%), compared to Core (6%) (Bertran-Gonzalez et al., 2008).

Like the core, the D1 and D2 containing MSNs have distinct projections. The D1 MSNs send a dense projection to the VTA (Lu et al., 1998; Zhou et al., 2003), while the lateral shell has reciprocal connections with the lateral VTA, and projects to the SNc (like the core). The medial shell has reciprocal connections with the medial VTA and projects to lateral VTA as well (Zhou et al., 2003; Ikemoto, 2007). The shell also projects to the ventral pallidum, maintaining a medial-lateral divide within shell to similar divisions in VP (Ikemoto, 2007).

Connectivity through the basal ganglia

Studies tracing the afferents from cortex to striatum suggest there may be five corticostriatal loops (Alexander et al., 1986). The main function of each of these loops is determined by the specific cortical input it receives. Each loop acts on a focused part of striatum, and causes inhibition to the corresponding output nuclei of the basal ganglia, in either the GPi or the SNr. This causes disinhibition of the thalamus and its corresponding projection back to the cortex (Chevalier and Deniau, 1990). A 'spiral' of successive projections from striatal regions to dopamine cells that then project to adjacent striatal regions is proposed to follow a shell-core-DMS-DLS sequence (Maurin et al., 1999; Haber et al., 2000).

Importantly, Joel and Weiner proposed a modification to the idea of parallel loops. Rather than strictly closed loops, they argued that there is interaction among different loops due to interconnections within channels (Joel and Weiner, 1997, 2000b)

(Joel and Weiner, 1994; Haber et al., 2000). Due to the interaction between the loops, it has been suggested that there is a functional hierarchy of striatal and cortical circuits (Redgrave et al., 1999). In addition, Haber proposed that connections between cortical areas that project to the striatal regions parallel the spiral, so that the shell projections to the cortex provide more feedforward inputs than the feedback inputs from the core projecting cortical region. These cortical projections also tend to innervate more than just their specific section of striatum, and instead extend to part of the adjacent striatal region, giving rise to the more integrative loop structure (Haber, 2003).

Connectivity throughout the cortex and basal ganglia determines the type of information that is transmitted. The sensorimotor loop refers to the connections between the DLS, the premotor cortex and the primary motor cortex. The main role of this loop is the selection of actions based on sensory and motor information, and is also implicated in the acquisition and expression of habitual behavior (Alexander et al., 1986; Yin and Knowlton, 2006; Packard and Knowlton, 2002; Featherstone and McDonald, 2004b). The associative loop refers to the connections between the DMS and the parietal cortex, the PFC and the frontal eye field (Cheatwood et al., 2003; Alexander et al., 1986). The main role of the associative loop is for orientation, attention and working memory, and has been implicated in learning and storing information relating actions to outcomes (Yin et al., 2005a; Yin and Knowlton, 2006).

Role of the basal ganglia circuit in behavior:

Nucleus Accumbens:

In general, manipulations to the NAc cause a change in spontaneous locomotion (Pennartz et al., 1994). In the core, blocking NMDA receptors leads to spatial deficits, such as not learning a path to rewards (Kelley, 1999), learning spatial sequences (Bauter et al., 2003) or finding platforms in a Morris water maze (Sargolini et al., 2003). Lesions between hippocampus and NAC lead to deficits in constructing paths, or forming new paths (Whishaw et al., 1995; Gorny et al., 2002).

Lesions of the shell do not produce spatial task deficits (Kelley et al., 2005; Kelley, 1999). Instead, Kelley and colleagues implicate the shell in free-feeding behavior, as well as approach behavior (Kelley et al., 2005), most likely through connections with the lateral hypothalamus (Kelley, 1999). Infusion of an AMPA receptor antagonist, or a GABA receptor agonist into the Shell stimulates feeding behavior (Kelley and Swanson, 1997; Basso and Kelley, 1999; Stratford and Kelley, 1997). Kelley also demonstrated that blockade of the Core impairs acquisition of a lever response task for food, but not on performance (Kelley and Swanson, 1997). This may have been due to a motivational deficit, which is supported by studies from Salamone, who show that dopamine depletion of the NAC causes rats to stop performing when more effort is required (Aberman and Salamone, 1999).

An alternative theory to the ventral striatum's role in goal-directed instrumental behavior is that it is only necessary for tasks that have long intervals between stimuli. Studies seeking to parse this out have shown that when animals are given the choice between a smaller reward given now, versus a larger reward after a delay, rats with lesions to the accumbens core make impulsive choices, and do not choose the larger, delayed reward (Cardinal et al., 2001). It has also been shown that in primates, the neural

activity in the VS across a delay does represent the expectation of reward (Schultz et al., 2000). Additionally, single unit studies in rats have shown that firing rates in the current trial are greatly affected by reward history (Kim et al., 2007), (Goldstein et al., 2012).

Recording studies in the NAc Core have found firing in anticipation of a variety of task related events, including cues that predict reward as well as reward receipt (Carelli and Deadwyler, 1994; Nicola et al., 2004a; Taha et al., 2007; Ito and Doya, 2009; Kim et al., 2009; Kimchi and Laubach, 2009b; van der Meer and Redish, 2009b; Goldstein et al., 2012). Since some of these studies have also shown neural firing patterns that anticipate the value of an outcome by firing more for the preferred reward, it is thought that the ventral striatum may help aid animals in making a decision by computing the relative value of a potential action. This role is further supported by the fact that few studies find neural firing in the NAC related to an upcoming choice prior to action selection, indicating that its role is in either computing value, or updating value after a choice is made and the outcome is revealed (Ito and Doya, 2009; Kim et al., 2009; Kimchi and Laubach, 2009b).

Evidence from Balleine and Killcross demonstrate that the accumbens is not required for goal directed behavior. Findings from lesion studies show that damage to the NAC does not affect sensitivity to reinforcer values (Balleine and Killcross, 1994). When they lesioned the NAC, they found that rats were still sensitive to changes in contingency, and that rats were still sensitive to the value of the outcome (Balleine and Killcross, 1994; Corbit et al., 2001).

The NAC has a role in motivational behavior. Cues that are associated with reward elicit attention and approach behavior from the rat, and are crucial to appetitive

and consummatory behavior. Damage to the NAC core during the acquisition or performance of Pavlovian tasks demonstrates its role for the ongoing maintenance of this behavior (Parkinson et al., 2000; Cardinal et al., 2002a). Another aspect of motivation can be measured as vigor, and studies that lesion the NAC core show a reduced rate of responding and abolishes the effect of Pavlovian instrumental transfer (Balleine and Killcross, 1994; Corbit and Balleine, 2011).

Another hypothesis for the role of the ventral striatum in task responding is the flexible approach hypothesis (Nicola, 2010). In this series of experiments, dopamine was depleted in the NAC core. The rat's main deficit was a reduction in operant responding, but this was due to the fact that animals had a harder time generating an approach behavior. Once the animals engaged in the start of the task, their performance was normal. The deficit was in beginning the task, especially if it required a flexible (ie, not stereotyped) behavior (Nicola, 2010). Single unit studies have found evidence that neurons encode information about this flexible approach behavior (McGinty et al., 2013).

The role of the Nucleus Accumbens Shell in decision-making is less known, although in some studies it appears to have overlap with the Core. Lesions to the Shell impairs specific responses in PIT (Corbit et al., 2001; Corbit and Balleine, 2011), while injection of amphetamine leads to an increase in response to levers associated with food (Parkinson et al., 1999). Somewhat confusingly, lesions of the Shell do not disrupt Pavlovian conditioning to aversive cues, and do not impair appetitive Pavlovian approach behavior (Parkinson et al., 2000; Parkinson et al., 1999). Dopamine release has been show in the Shell when exposed to unconditioned stimuli (or primary reinforcers) that are unexpected (Ito et al., 2000; Bassareo and Di Chiara, 1999). However, conditioned

stimuli do not elevate dopamine levels in the Shell, but do instead elevate these levels in the Core (Ito et al., 2000; Bassareo and Di Chiara, 1999; Wilkinson et al., 1998; Roitman et al., 2004; Day et al., 2007; Jones et al., 2010; Cacciapaglia et al., 2012; Cacciapaglia et al., 2011).

Dorsal Striatum:

Rats with lesions to the Dorsal Lateral Striatum (DLS) are impaired in simple discrimination tasks using different cues and responses (Packard et al., 1989; Reading et al., 1991; McDonald and White, 1993; McDonald and Hong, 2004). Tasks that require the acquisition of a stimulus-response association, such as pressing a lever or pulling a chain when a specific cue (such as a light) comes on in order to receive a reward are disrupted by lesions of the DLS- both in the acquisition and performance of the task (Featherstone and McDonald, 2004b).

Additionally, it is thought that stimulus-response representations are both encoded and potentially stored in the DLS. When rats were taught to lever press for a food reward (sucrose), and then tested in a devaluation procedure, where the sucrose was paired with feeling sick, normal rats and rats with DLS lesions both showed a reduced response to sucrose. However, when rats with DLS lesions were then tested in an instrumental task, where lever press used to be paired with reward (but this is done in extinction, so no reward is present), they reduce responding to the lever when normal rats continue to press the lever. In this scenario, it is proposed that the normal rats are responding to the lever in a habitual manner, and are using the DLS, while the DLS-lesioned rats have not acquired

a the stimulus-response (SR) association, and are instead using the DMS, which retains sensitivity to reinforcer devaluation (Yin et al., 2006; Yin et al., 2004).

Reports from the primate literature have found single unit examples of cue-responsive neurons in the DLS (Apicella et al., 1992; Vicente et al., 2012; Yamada et al., 2004; Hori et al., 2009). From the hypothesis that the DLS encodes information related to stimulus-response, it would make sense to find single-unit firing of the cue, or stimulus, in the DLS. This prediction has not held up in many rodent single unit studies, showing very little, if any, cue-specific firing in the rodent DLS (Gage et al., 2010; Berke et al., 2009; Thorn et al., 2010).

The Dorsal Medial Striatum (DMS) has a broad role in a range of behaviors. There have been neural correlates of direction, location and movement selectivity in the DMS (Wiener, 1993; Kim et al., 2009; Mizumori et al., 2009). However, the role of the DMS goes beyond just spatial navigation, and includes a role in behavioral flexibility (Ragozzino et al., 2002), and includes encoding stimuli and actions in tasks that go beyond allocentric encoding (Ito and Doya, 2009; Kimchi and Laubach, 2009b).

In instrumental learning tasks, the DMS is thought to play a role in response-outcome learning. When testing rats on a lever press task for food reward, increased training renders the behavior more habitual, and less susceptible to reinforcer devaluation procedures. If the DMS is lesioned, the rat no longer forms response-outcome associations, which is the hallmark of flexible, goal-directed behavior (Yin and Knowlton, 2004; Yin et al., 2005a; Yin and Knowlton, 2006; Daw et al., 2005).

Reinforcement Learning and the Basal Ganglia

Reinforcement Learning is a computational approach to understand how learning different types of action associations maximizes the overall accumulation of rewards (or appetitive outcomes) (Sutton and Barto 1998). Using terms from reinforcement learning theory, an individual ('agent') can represent the world as a set of states, of which the agent can take one of a set of potential actions. A state (or state space) is something like a location in a maze, or the presence or absence of a stimulus in an operant box (Dayan and Niv, 2008). An action in a specific state leads to the next state, as well as a possible reward (Sutton ad Barto 1998).

Actor-Critic learning

One approach to understanding the break down of striatal components of action selection that comes from Reinforcement Learning theory is called 'Actor-Critic' learning. In seeking to maximize rewards and minimize punishment, the agent has to learn from experience. To do so, the agent relies on a reward prediction error (RPE) that comes from the difference of the expected outcome of a choice vs. the actual outcome. In the actor-critic view of decision-making, there is a component that learns the stimulus-action policies (actor), while another area learns to predict rewards (critic) (Barto, 1995). Evidence from rodent lesion studies demonstrates a role for the dorsal striatum as actor and ventral striatum as the critic (Packard and Knowlton, 2002; Cardinal et al., 2002a). Corresponding human imaging studies add to these results (O'Doherty et al., 2004; Tricomi et al., 2004; Tricomi et al., 2009). Signals from the dopaminergic neurons in the SNC to the dorsal striatum are thought to influence synaptic plasticity, which in turn, guides learning of the action-selection policy, as the actor (Joel et al., 2002).

Modeling decision-making through the actor-critic framework has the added bonus of fitting in nicely with other classical psychological distinctions: the Pavlovian system and the Instrumental system. With Pavlovian learning, the agent learns the stimulus-outcome relationship (critic), while in instrumental learning the agent learns which actions are advantageous (actor). Within instrumental learning, a further breakdown can be seen in how the animal is actually responding- if it is habitual or goal-directed. When behavior is very well learned, and is no longer sensitive to changes in the value of the outcome, it is considered habitual, whereas behavior that is still sensitive to contingency degradation is considered goal-directed.

Due to anatomical connections between the motor and limbic system (Mogenson et al., 1980), as well as connections to the dopamine neurons in the VTA (Haber and Brucker, 2009), the nucleus accumbens has been thought of as the 'critic' in this framework (Joel et al., 2002). Studies using fast scan cyclic voltammetry (FSCV) have found value-predicting error signals in dopamine levels in the NAc (Cheer et al., 2007; Day et al., 2007). Studies examining the role of the NAc in response to receiving a reward has found reward-related firing that occurs shortly after an animal is rewarded (Lavoie and Mizumori, 1994; Taha and Fields, 2005), as well as a ramp-like increase in firing as an animal approaches reward (Lavoie and Mizumori, 1994; van der Meer and Redish, 2011). However, single unit studies in the rodent have looked for, but not found, encoding of an RPE signal in NAc spiking activity (Roesch et al., 2009; Ito and Doya, 2009).

The actor-critic framework relies on stimulus-response learning only, and cannot explain aspects of goal-directed behavior like devaluation sensitivity. The most important

distinction is that in the actor-critic form of learning and decision-making, the agent must experience a particular reward in a certain state to understand that it is no longer valued. Contrary to this, tests using satiety or devaluation have found that even on the first trial, before experiencing the reward, a rat can already have a representation of that goal in mind, and act accordingly. This means that he is not using newly learned action-value representations to modify the reward value in a new sated or devalued state. A new explanation of behavior using reinforcement learning was needed.

Model-free vs. Model-based encoding

Due to the difficulty in explaining how an animal could change his response to a devalued outcome, without having experienced that 'state', theorists have more recently begun using another set of ideas from machine learning to understand how each region is representing information. An important component for decision-making in the RL framework is the use of 'action-values'. An action-value is an estimate on how much future reward is expected by taking a particular action in a given state.

There are two main differences in how an action is selected, based on how the agent represents information. In a model-free representation of these states, each potential action has a cached 'action-value', where the action itself is given a reward prediction error (i.e., the likelihood of reward), and an action is chosen based on that representation-not of the goal itself. Using a model-free representation, action selection is due to two interacting components. One part learns to predict a reward value while another part forms the rule about what action to select when in a particular state (Daw et al., 2005; Rangel et al., 2008; Redish et al., 2008). This form of action-selection is faster and less

computationally intense, since the animal uses a cached value to determine the potential value of an action, without representing the actual outcome itself.

Using a model-based representation, the agent learns the value of different actions through experience and constructs a model of his environment. When making a decision, the agent utilizes information about the current state, and chooses an action based on an estimation of the goal, by searching through what the consequences of each action lead to. The agent forms a model of the world in the current state, and uses an on-line search process to determine the potential rewards available depending on the state. The agent then thinks through the consequences of potential actions in order to determine which action to take (Daw et al., 2005; Rangel et al., 2008; Redish et al., 2008). A representation of real-world model-free and model-based encoding is shown in Figure 2C.

An important prediction from the model-based and model-free decision-making framework is that some areas of the brain should represent action-values. In primates, this encoding has been found in the dorsal striatum, as well as the GPi and the supplemental motor area (Samejima et al., 2005; Pasquereau et al., 2007; Wunderlich et al., 2009; Hori et al., 2009; Lau and Glimcher, 2008). In rodents, this signal is not as robust, but significant proportions have been found in the dorsal and ventral striatum (Ito and Doya, 2009; Kim et al., 2009; Roesch et al., 2009). However other studies examining single units in the ventral striatum of rodents has not found any integration of action and value (Kimchi and Laubach, 2009b).

Just as action-outcome has been mapped onto the DMS, it is also proposed as the site responsible for model-based encoding. The DLS has been proposed to carry model-

free information, similar to its role in processing habits and stimulus-response associations (Daw et al., 2005). Model-based, or goal directed behavior is flexible, and can quickly adapt to changes in reward contingency, but is computationally expensive. Model-free, or stimulus-response/habitual behavior is more efficient, and typically involves less processing time, but is inflexible.

When making a decision, information about the value or consequence of potential actions must be transmitted to a place where action-selection occurs. Previous work trying to determine the exact locus of action-selection may depend on the nature of the behavioral task, and action-selection regions have been discovered in a variety of regions. Behaviors that involve fixed stimulus-action associations, or highly learned motor sequences appear to rely on the dorsolateral striatum (Knowlton et al., 1996; Hikosaka et al., 1999; Yin et al., 2006). However in studies that involve more flexible decision-making, such as a dynamic foraging task, the signal related to an animal's upcoming choice has appeared earliest in the medial motor cortex (Sul et al., 2011). The representation of the upcoming action has been found in the dorsal striatum in both primates and rodents (Samejima et al., 2005; Pasquereau et al., 2007; Kim et al., 2009; Pasupathy and Miller, 2005). Representation of the upcoming action in the NAc has also been found (Roesch et al., 2009), as well as not found (Ito and Doya, 2009; Kim et al., 2009; Kim et al., 2007), leading to open questions about where the actual action-selection decision occurs.

Another important aspect to the decision-making process is the ability of the agent to understand the relationship between an action that was just performed, and its consequences, as well as the influence of that action on future behavior. For example, in a

game of chess, it is important to be able to understand that the move just made was either good or bad. This is often referred to as the temporal credit assignment problem (Sutton and Barto 1998). Brain regions that are important in an animals' ability to correctly alter decisions based on an outcome include the OFC (Schoenbaum et al., 2002), the prefrontal and posterior parietal cortex of primates (Barraclough et al., 2004; Seo and Lee, 2009), and the frontal cortex and striatum of rodents and primates (Kim et al., 2007; Kim et al., 2009; Sul et al., 2010; Sul et al., 2011; Nakamura et al., 2012).

In order to update value functions in a model of reinforcement learning, a reward prediction error is used. This is defined as the difference between the actual reward and the reward expected by the current value function (Sutton and Barto 1998). These signals were first isolated in putative midbrain dopamine neurons (Schultz, 2006), but have also been found in other regions including the globus pallidus (Hong and Hikosaka, 2008), orbitofrontal cortex (Sul et al., 2010) and striatum (Kim et al., 2009; Oyama et al., 2010; Stalnaker et al., 2012).

Early studies of reward originally thought of dopamine as the reward signal in the brain (Wise, 1978). Studies in primates done in the '90s found that the role of the dopaminergic neurons was more complex than simply a reward signal. In experiments done in the Schultz lab, monkeys were given rewarding stimuli, such as food and water, and recordings were done during these instrumental or Pavlovian conditioning tasks. During learning, the recorded cells did show phasic bursts of firing when given a reward, but if the reward was consistently preceded by a cue, such as a tone, the dopaminergic response to the reward gradually decreased. This contradicts the theory that dopamine always equals reward. Upon further examination of the monkey's behavior and the

subsequent neural firing rates, they observed that the subjects were exhibiting anticipatory behavior to the cues that predicted reward, and the neural firing rates followed (Ljungberg et al., 1992; Romo and Schultz, 1990; Schultz et al., 1993).

When computational theorists began applying reinforcement learning algorithms to neuroscience, this dopaminergic signal was obvious- it was a reward prediction error; (Montague, 1993, 1995; Montague and Sejnowski, 1994). The key was that the RPE signaled an unexpected reward. In initial learning, the primates did not expect reward delivery, so the dopaminergic neurons fired when reward appeared. With learning, they began to expect the reward, so the neurons gradually stopped firing. However, the cue that predicted the reward became the unexpected event, and had reward predictive power, hence eliciting firing from the dopaminergic neurons (Takikawa et al., 2004; Hollerman et al., 1998).

Verification of this theory has been demonstrated in a variety of tasks (Bayer and Glimcher, 2005; Hollerman et al., 1998). In fact, when using probabilistic rewarding tasks, or tasks where different reward magnitudes are predicted, and the relationship between cue and reward delivery depends on a reward probability or magnitude, rather than a simple 1:1 reward ratio, the conditioned stimuli that predicts the reward elicits a phasic dopaminergic response that is proportional to the probability/magnitude of expected reward (Fiorillo et al., 2003; Tobler et al., 2005). In rodents, the relationship between the conditioned stimulus and reward has been shown with phasic levels of dopamine (Day et al., 2007; Walton et al., 2006).

While in the decision-making process, the value of the action that was actually chosen must also be stored in order to assess the outcome of choosing that action, and is

often referred to as the chosen value. Various brain regions carrying this signal include the medial frontal cortex, the orbitofrontal cortex and the striatum (Padoa-Schioppa and Assad, 2006; Lau and Glimcher, 2008; Kim et al., 2009; Sul et al., 2010; Cai et al., 2011). This signal, paired with the RPE, in some RL theories, can then inform the rat about the next state, and the updated action-value associations.

Contrary to the idea that the phasic burst of dopamine acts as an RPE in reinforcement learning theory, Berridge (Berridge, 2007) argues that phasic dopamine is not crucial to learning, but is instead a signal of incentive salience, which acts to maintain or repeat whatever current action is currently rewarding. The dopamine signal is therefore a signal of 'wanting'.

Mapping onto strict subregions- with difficulty:

A series of lesion studies came up with a potential explanation for how there could be both stimulus-response and goal directed behavior simultaneously. In these studies, it is proposed that different subregions of the striatum represent separate aspects of instrumental decision-making. Pharmacological blocking of the DMS causes goal-directed actions to become habitual, while on the other hand, blocking activity in the DLS causes previously habitual actions to become goal-directed (Yin et al., 2006; Yin et al., 2005a; Yin and Knowlton, 2004; Balleine et al., 2009). This anatomical separation keeps the actor-critic RL theory, as the model-free representation in the DLS and Core, while a model-based, or goal-directed representation also resides in the DMS and Shell (Figure 2B).

These theoretical frameworks have held up under some lesion and imaging studies, but there are other aspects that question the overall architecture. The first issue is that the ventral striatum, which is supposed to function as a critic in the typical RL theory, shows model-based tendencies, such as sensitivity to devaluation (Van Der Meer and Redish, 2009b; van der Meer et al., 2010). Nicola (2010) has found that the NAC is involved when a rat can flexibly approach an aspect of reward, but if the task involves stereotyped, or habitual actions, the ventral striatum is less involved, again going against a role in habits or model-free learning (Nicola, 2010). Studies seeking to find precise firing rate changes in aspects of the dorsal striatum that track onto the model-free vs. model-based framework are often conflicting, showing similar patterns of activity related to stimulus-response and response-outcome encoding in both DMS and DLS (Stalnaker et al., 2010; Thorn et al., 2010; Kim et al., 2013). A possible explanation suggests that these hypothetical systems are more interactive, which is consistent with an overlapping and interconnected loop representation of information in the basal ganglia, and tracks onto anatomical studies of interconnection between regions (Alexander et al., 1986; Joel and Weiner, 2000).

Single unit studies examining the role of the dorsal striatum in habitual actions have found a distinction between DMS and DLS at the beginning and end of action sequences (Barnes et al., 2005; Thorn et al., 2010), as well as task-related firing changes that develop with experience (Barnes et al., 2005; van der Meer et al., 2010; Thorn et al., 2010). However, distinct separations based on anatomical location into stimulus-response vs. action-outcome representation leads to mixed results (Kimchi and Laubach, 2009b; Stalnaker et al., 2010; (Thorn et al., 2010; Kim et al., 2013). Kim (2013) found overlap

between the DMS and DLS in the animal's goal choice, the outcome, as well as an action value signal before the choice was revealed (Kim 2013). Stalnaker 2010 (Stalnaker et al., 2010) found significant overlap between stimulus-response encoding and response-outcome encoding in both the DMS and DLS, with no modification in firing rate due to value. There is clearly a gap in the understanding of the role of the striatum as it relates to these elegant models of behavior when examining studies done using lesions vs. recording single units.

The biggest difficulty in finding the neural correlates of model-free vs. model-based behavior appear mostly in finding a specific underpinning for the model-free neural firing rates. In most studies that seek to compare contrasting strategies, areas that had previously been thought to only be involved in model-free behavior show model-based differentiation. In the human fMRI literature, there are two main types of tasks that can be broadly broken down: sequential decision-making and tasks involving the inference or observation in a change in reward value. In all of these studies, all areas of the brain are activated anywhere reward information is processed, even if these tasks are not explicitly using model-free information (Glascher et al., 2010; Daw et al., 2011; van der Meer et al., 2010; van der Meer and Redish, 2010; Wunderlich et al., 2012; Abe and Lee, 2011). Most importantly, an area of the brain thought to be used only in the model-free system, the human ventral striatum, has shown RPE correlates in a model-based task (Daw et al., 2011; Simon and Daw, 2011; Wimmer et al., 2012).

Orbitofrontal Cortex

A final brain region that is relevant to a study on decision-making and reward is the orbitofrontal cortex. Besides the striatum, reward prediction and movement-related neurons have been isolated in the orbitofrontal cortex (OFC). Convergent evidence from humans, non-human primates and rats implicates the OFC in a variety of behaviors relating to decision-making, flexible actions and reward (Rolls, 2000). Disorders involving the OFC include depression and obsessive-compulsive disorder. Understanding how the OFC acts in concert with the striatum, in regards to updating information on internal states is a crucial piece of information missing in the literature.

Anatomy: The orbitofrontal cortex in rodents receives inputs from other areas of the cortex, including areas that receive sensory information (Price 1985, Price 2007, Dalley et al., 2004; Schoenbaum and Roesch, 2005; Schoenbaum and Esber, 2010), as well as spatiomotor information, including the posterior parietal cortex and medial agranular cortex (Reep 1996). The more medial areas of the rat cortex receive stronger afferents from the limbic system, including the amygdala and hippocampus (Ongur and Price, 2000; Price, 2007).

The ventro-lateral area of the rodent orbital cortex projects more to sensorimotor structures (Coffield et al., 1992, Reep et al., 1994, 1996). Additionally, the distribution of projections to the dorsal striatum shows that more medial orbital regions target the medial striatum, while the VLO innervates the central regions and the more lateral orbital cortex projects to the more lateral areas of the striatum. The connections of the OFC also make it a unique target to determine if sensory properties are more explicitly encoded in the OFC, compared to striatum.

Orbitofrontal Cortex and Behavior:

A prominent area of study demonstrating a role for the rodent OFC in outcome-guided behavior comes from reversal studies. In reversal studies, a rodent first learns to discriminate between two cues, one that predicts a reward, and another that predicts either an aversive event, or no outcome. Once this is learned, the contingencies are changed, so that the previously rewarded cue now predicts the opposite outcome (either the aversive event or nothing), while the other cue now leads to reward. Learning is achieved when the animal demonstrates that he has learned the new cue-outcome pairing. Studies using lesions in the OFC have shown that it is necessary in order for rapid reversal learning to occur (Schoenbaum et al., 2002; Izquierdo et al., 2004). One specific aspect of these results that is crucial to impart is that the OFC is not necessary for the initial discrimination learning- it is only when the contingencies are changed that the OFC becomes necessary.

A variety of theories try to explain what, specifically, the OFC is doing in these reversal studies. Some think that the role of the OFC is to inhibit a previously learned response (Eagle et al., 2008; Izquierdo and Jentsch, 2012). Other theories state that the OFC is using and/or assigning value to the cue, in a common currency (Levy and Glimcher, 2011; Padoa-Schioppa and Cai, 2011; Padoa-Schioppa and Assad, 2006). Additionally, other theories think that the OFC is necessary for predicting and integrating specific features of an upcoming event, possibly without actually signaling value (Clark et al., 2012; Delamater, 2007; McDannald et al., 2012; Schoenbaum et al., 2009).

Another area of study demonstrating a clear role of the OFC in behavior is by utilizing similar procedures discussed above- outcome devaluation. In these studies, an

animal learns that a cue predicts a specific outcome. After some amount of learning, the outcome is devalued, either through pairing it with feeling sick, or feeding an animal to satiety. Lesions of the OFC in rodents show a deficit in performance when testing an animal with outcome devaluation (Gallagher et al., 1999; Pickens et al., 2005; Pickens et al., 2003). These deficits are not due to learning- the animals still acquire the Pavlovian or instrumental response. The OFC is also not necessary when choosing between rewards when they are actually present (Gallagher et al., 1999; Pickens et al., 2005; Pickens et al., 2003). The specific deficit is seen when the animal needs to integrate information about the reward, and pair that with a predictive cue to update the new value of the reward.

Additional tests used to probe the specific aspect of outcome encoding seen in the OFC have utilized two more behavioral tasks. In one procedure, called 'unblocking', a rat learns the relationship between a cue and an outcome. An additional cue is then added to the first cue, while also increasing the size of the reward. The new cue has new information, because it is adding information about the increased value of the reward- not the specific identity of the outcome. The response to the new cue is similar to the second behavioral task used to probe OFC function, called Pavlovian-to-instrumental transfer (PIT). In these studies, rodents learn that two separate responses lead to two separate, but valued rewards. Animals also learn that two different cues lead to these different rewards. Once the animal learns both the Pavlovian and instrumental response, a test occurs where the cue is given in the presence of the instrumental response, with no reward. During the presentation of the specific Pavlovian learned cue, instrumental actions related to the prediction of that cue are typically enhanced.

The contribution of the OFC to both of these behavioral tasks appears to support a

role in specific outcome identity. When lesions are made prior to training, an unblocking procedure based on value remains intact, while another based on reward features is impaired (McDannald et al., 2011; Burke et al., 2008). Studies utilizing PIT and OFC lesions show that there is no deficit in the acquisition of the Pavlovian association or the instrumental learning, but OFC lesions do impair the performance of specific transfer (Ostlund and Balleine, 2007; Scarlet et al., 2012).

Single unit studies paint a more complicated picture. One difficulty in understanding the function of the OFC is the species differences between primate and rodent studies. In both species, the OFC appears to learn the value of the stimulus (Burke et al., 2008), and is used to make adaptive decisions (Rudebeck et al., 2006). Additionally, both species encode decision related information (Schoenbaum et al., 1998; (Roesch et al., 2006; Kepecs et al., 2008; van Duuren et al., 2007; van Duuren et al., 2009). One striking difference is that rodent studies have found response (or action) specific encoding in the OFC (Furuyashiki et al., 2008; Feierstein et al., 2006), while primate studies do not (Tremblay and Schultz, 1999; Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006).

Besides response differences, another species difference seen in recording studies is the integration of value signals. As discussed before, rodent lesion and some recording studies have found that the OFC encodes specific information about the outcome, but is not necessarily doing so by integrating an abstract value signal (Roesch et al., 2006; Burke et al., 2008). In contrast, most primate research finds a signal related to abstract value (Padoa-Schioppa and Assad, 2006; Morrison and Salzman, 2009; Kennerley and Wallis, 2009b).

Because of the findings that movement is not encoded in primate OFC, and that there is an integration of multiple signals for value, primate researchers have proposed an alternate theory than that discussed above by Balleine and Schoenbaum. Specifically, studies comparing neural firing rates in primates have demonstrated that the OFC contributes to encoding value in a common currency (Padoa-Schioppa and Assad, 2006; (Padoa-Schioppa and Cai, 2011). In this way, different goods can be compared using a common neural substrate- there is a way to compare apples to oranges (Padoa-Schioppa and Cai, 2011). Additionally, single unit studies have found neural correlates in the OFC for reward probability (Kennerley and Wallis, 2009b), response requirement and reward identity (Rolls and Baylis, 1994).

There have not been as many single unit studies in the OFC of rodents studying different reward types, so most single unit studies find neural firing rates that may encode an aspect of value in upcoming reward (van Duuren et al., 2007; van Duuren et al., 2009; Feierstein et al., 2006; Roesch et al., 2006). However, there does appear to be some identity encoding, regardless of preference, in the rodent OFC as well as specific action-related activity (Furuyashiki et al., 2008; Feierstein et al., 2006; Young and Shapiro, 2011; Roesch et al., 2006) that is not seen in primate OFC (Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006; Abe and Lee, 2011), and leads to confusion about the role of the rodent OFC in encoding a 'common currency'. Compounding that difficulty is that more recent primate work has found evidence of specific direction-related coding in the OFC, when a task requiring more movement was used (Luk and Wallis, 2013; Tsujimoto et al., 2009).

Internal States:

There are few studies that directly manipulate internal motivational states, while directly testing decision-making processes. From these studies, evidence exists that internal state is encoded in the basal ganglia, and may be a crucial component of reward-based neuronal firing rates. The Berridge lab demonstrated how shifts in internal representation of rewards influence neurons in the ventral pallidum. These studies utilized a taste reactivity paradigm, along with Pavlovian conditioning, to pair a stimulus with a tastant, delivered directly into the oral cavity (Tindell et al., 2004; Tindell et al., 2006). In these studies, they manipulated the rat's internal salt content to change a previously aversive salt taste into a rewarding stimulus under salt deprivation. This distinct change in behavioral outcome was mirrored by a change in firing rate in the ventral pallidum. Not only did the firing rates change when the salt tastant was orally infused, but the firing patterns also changed significantly to the cue predicting the infusion, indicating that the rat knew the salt would be rewarding in his salt-depleted state. These studies demonstrate a specific subset of neurons responsible for tracking the internal state of an animal in the basal ganglia (Smith et al., 2009).

This evidence indicates that specific cells can modulate their firing rates for changing reward values, as well as modifying firing rates to the cues that predict the rewards. The Carelli lab compared natural rewards to drug reinforcers and found that cocaine self-administration elicits responding in specific neurons that are in a separate population compared to neurons that track food and water reward (Carelli, 2004; Carelli and Deadwyler, 1994; Carelli et al., 2000). They reported on a portion of the neurons that overlapped, and coded both food and water reward (68%) in the ventral striatum. The

remaining neurons were either selective for one reward or the other, or indistinguishable, but since the original hypothesis was testing natural vs. drug reward, food and water specific neurons were not included. Between the Berridge and Carelli studies, there is ample evidence to suggest that internal states are 'online' while a rat is performing a task. Manipulations of internal states can be done in animals, by regulating access to food and water, or by salt depletion, which consequently leads to different firing patterns (Rolls, 1989).

Additional evidence for internally motivated neural firing comes from a study done by the Shapiro lab while recording from the hippocampus. Rats were either food or water deprived, and had to make decisions based on their internal state in a contextual-based memory task, which the researchers then compared to the rat's behavior in a random foraging task (Kennedy and Shapiro, 2009). Importantly, they found that motivational state influenced hippocampal encoding during the contextual memory task, when the motivational cues were crucial to selecting a remembered, goal-based action. During the foraging task, when memory based cues were not involved, these same neurons did not determine behavior (i.e., they did not fire more for food or water). Kennedy and Shapiro demonstrated that internal motivation not only guides memory-based tasks, but these cues also reflect concurrent changes in firing rates. Their study did not compare food and water within the same trial, however.

OFC and internal state: Evidence also exists that the OFC encodes signals relating to hunger, and may in fact encode motivation. De Araujo et al. studied ensembles in the OFC, among other places while rats were fed to satiety. They found that OFC neurons encode a signal that reflected the animal's motivational state across the entire

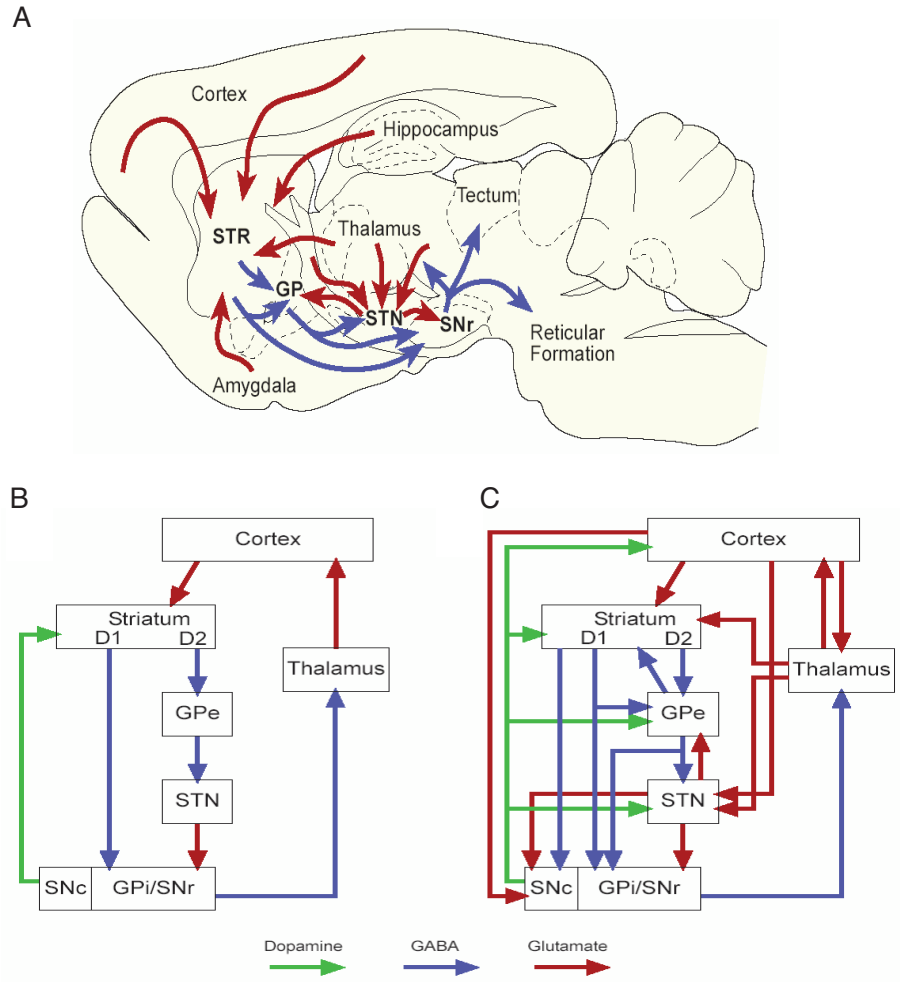
feeding cycle, with higher activity levels during the hunger phases (de Araujo et al., 2006). This is similar to early electrophysiology studies by Rolls et al that found a decrease in response to cues predicting food when the animal was sated (Rolls, 1989). There is evidence that neuronal populations in the OFC respond to separate food/water stimuli, but whether this is simply encoding a sensory property or it's motivational salience is unknown (Rolls, 2004).

The behavioral task used in this project presents a unique opportunity to test the relevant theories related to OFC function. This task disentangles reward from movement, so a definitive role for the OFC in encoding movement, absent reward can be tested. Additionally, the anatomical connections with sensory areas between OFC and other cortical regions, as well as downstream targets put it in a unique location to integrate sensory information with internal information in order to calculate value. On the other hand, the OFC may just encode the sensory properties of the reward identity itself, absent value, allowing us to answer the question: does the OFC encode specific reward identity, or is it encoding value? And how is action represented in the OFC, independent of reward?

The Carelli et al studies demonstrate that internal motivation may be encoded in the ventral striatum, while the Berridge et al and Shapiro et al studies show that these firing rate changes influence behavior and neuronal firing in other brain areas, indicating that the entire cortical-basal ganglia circuit may be involved in updating online information about internal motivation. The advantage of the behavioral task used in this study is the ability to distinctly dissociate internal motivation from specific movements, all while recording from multiple brain areas to solidify theories of how motivation,

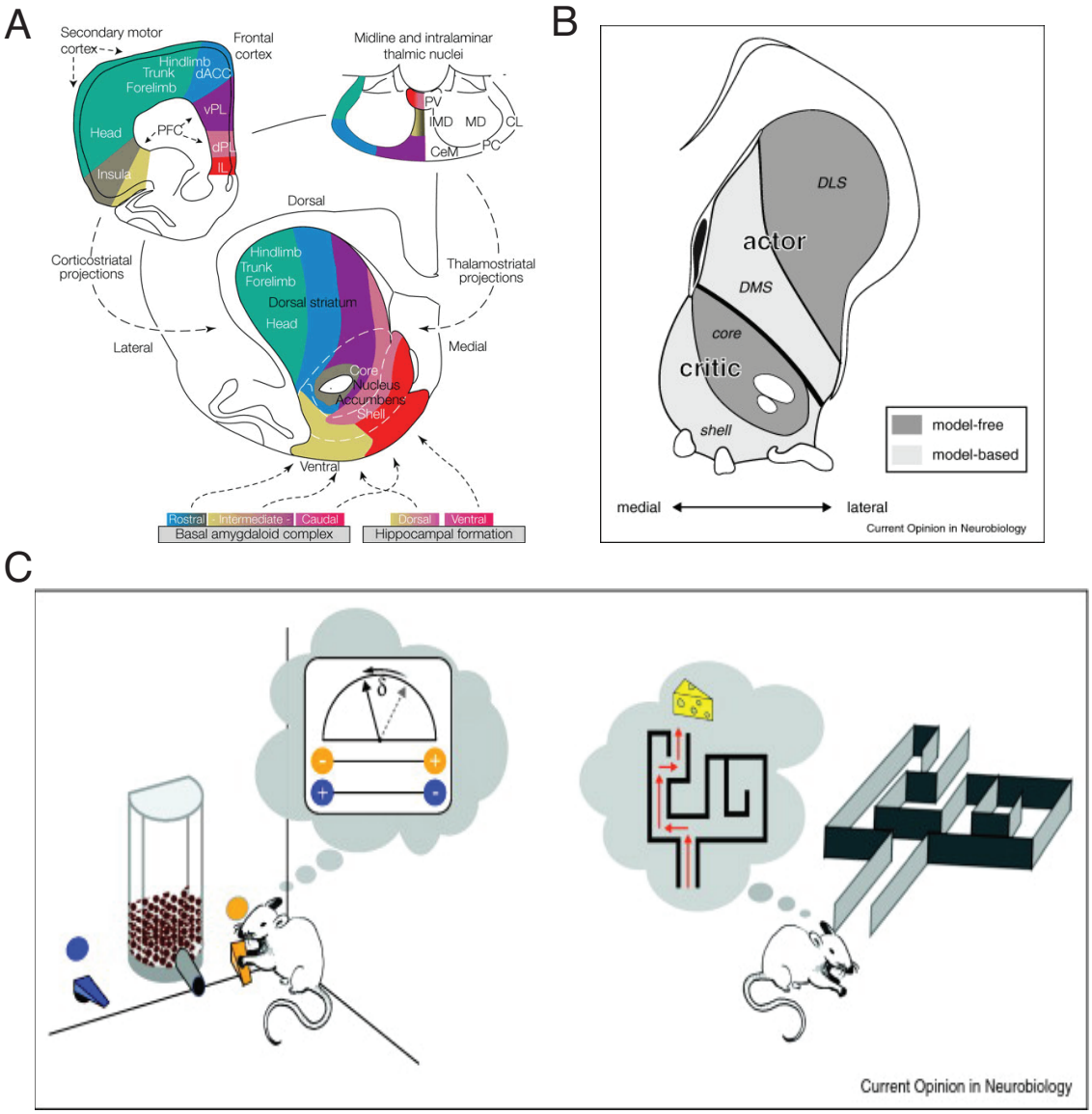
action and value are updated in the basal ganglia. Manipulation of internal state, combined with trials of flexible choice, allows us to determine how the neural firing rates correspond to behavior during an actual decision-making event that shows sensitivity to internal influences.

Figure 1: Basal Ganglia Anatomy



A: Rodent Basal Ganglia. The main input nuclei are the Striatum and the Subthalamic Nucleus (STN), which receive input from the thalamus, cortex, amygdala and hippocampus. The main output nuclei are the substrantia nigra, and the globus pallidus. Red arrows are excitatory connections while blue arrows denote inhibitory connections. (Redgrave, Scholarpedia, 2007).
B: Original Box and Arrow diagram from Albin et al. (1989). The direct pathway promotes behavior through D1 containing neurons in the striatum, and the indirect pathway suppresses behavior through D2-containing neurons (Redgrave, Scholarpedia, 2007).
C: The updated, more complicated organization of the Direct and Indirect pathways (Redgrave, Scholarpedia, 2007).

Figure 2: Reinforcement Learning and the Basal Ganglia



A: The topographical arrangement of cortical and thalamic inputs to the striatum show a distribution of dorsomedial to ventrolateral zones. Frontal cortical areas and their corresponding striatal projection zones are shown in the same colors. Abbreviations are listed in the Abbreviation section (Voorn et al., *Trends in Neuroscience*, 2004)

B: The proposed dorsal/ventral divide of the actor/critic framework, with further subdivision based on model-free and model-based Reinforcement Learning. Model-Free is shown in dark grey, and Model-Based is shown in light grey (Bornstein and Daw, *Current opinion in neurobiology*, 2011).

C: Demonstration of model-based and model-free reinforcement learning. In Model-Based computation, a mental map is used that has been learned through prior experience. This forward model utilizes on online search process to predict probabilities of upcoming reward, based on the available action options. In contrast, Model-Free action selection is based on learning the long-run values of specific actions, without having to build a map or model of the current environment (Dayan and Niv, *Current opinion in neurobiology*, 2008).

Chapter 2: Materials and Methods

Behavioral Task

All animal procedures were approved by the University of Michigan Committee on Animal Use and Care. Animals were housed on a 12:12 reversed light/dark cycle, with experiments performed during the dark phase. Long Evans rats (350-550g) were trained and tested using a modified reaction time task in a five nose-port operant chamber (Med Associates Inc, St Albans, VT). The operant box consisted of stainless steel grid floors with five nosepokes, a speaker, a video camera and a reward dispenser that physically separated the food and water ports. Infrared photobeam detectors recorded entry into every nosepoke and food or water receptacle entry.

Task Description: A trial began with a 500ms tone (1kHz = water, 4kHz = food), informing the rat which reward would be available following a correct performance. Immediately following the tone offset, one of the three central nosepokes was illuminated. The rat had to poke the lit port, prompting the illumination of an adjacent nosepoke (to his left or right, pseudorandomly). After correctly poking this second port, a white noise burst signaled to the rat that he was correct, and he could retrieve a reward. Food and water ports were physically separated, and each had a photo-beam sensor to monitor entry and exit, therefore reward was not actually delivered until the rat entered

the correct receptacle. Precise measurements of all actions were recorded using Labview Software and photobeam detectors on each nosepoke and food/water receptacles. Upon completion of a correctly rewarded trial, the rat had an 8-10 second inter-trial interval before a tone played to begin the next trial. The sequence of task events is illustrated in Figure 3A.

Behavioral Events: The task included 6 main behavioral events. The first event is 'Tone', which is centered on when the food/water tone began playing. The next event is 'Nose Center In', aligned to the time when the rat poked his nose into the first lit nose port. The task was self-paced, so there was no time limit between 'Tone' and 'Nose Center In'. The time window between these two events is considered the Initiation Time. The next event is 'Nose Center Out', which is aligned to the time the rat removed his nose from the first lit nose port. 'Nose Side In' was the next event, triggered by the rat entering the adjacent lit nosepoke. The time between 'Nose Center In' and 'Nose Side In' is considered the Movement Time. 'Nose Side Out' is aligned to when the rat exited the second nosepoke. 'Reward Port In' is aligned to the time when the rat entered the reward receptacle (before receiving the reward). The time from 'Nose Side Out' to 'Reward Port In' is considered the Reward Retrieval Time. Behavioral events are outlined in Figure 3A.

Task Training: Rats were initially trained to associate the 500ms tone (1kHz = water, 4kHz = food) with the appropriate reward, which was delivered in the separated food or water receptacle. After rats reliably learned the tone-reward pairings (correct reward

entry >80% of trials), they were then trained to poke a lit nosepoke before being presented with the food or water tone. The next level of training presented the rats with the tone first, then one lit nosepoke, which required the rats to remember the correct tone and enter the appropriate reward receptacle after poking 1 lit nosepoke. Once the rats achieved 80% reliability on the 1-poke version of the task, an additional nosepoke was introduced after the first correct poke, forcing the rats to move to the adjacent left or right nosepoke before entering the appropriate reward receptacle. All rats were trained using the same paradigm of increasing difficulty, culminating in the final behavioral task (outlined above). Training typically lasted 2-3 months. Once a rat achieved 80% accuracy while both food and water restricted on 3 consecutive days, he underwent drive implantation surgery.

Errors: Inappropriate actions during trial performance resulted in a 15 second trial timeout, signaled by illumination of the houselight. Procedural errors include instances when the rat poked an inappropriate nosepoke (one that is not lit) at any point after the start of a trial. Reward port errors were when the rat went to the inappropriate reward receptacle (ex. going to the food port on a water cued trial). Procedural errors are not included in the behavioral analysis. If the rat performed either a procedural or reward error, after the trial time out, the rat had to repeat the same trial type until correct. This ensured that the rat completed both left and right movements, and prevented the rat from perseverating on only one reward receptacle.

Free choice: On select trials (20-30% of total trials), the rat heard a 12 kHz tone lasting 50ms. On these trials, once the rat correctly poked the 2 nosepokes, he was allowed to choose either reward by poking his nose in the desired reward receptacle. Whichever port he entered first triggered reward delivery. These trials allowed us to probe which of the 2 rewards were considered more desirable to the rat (see below).

Internal Motivation Manipulation: Before each testing session rats were either food-restricted, water-restricted, or both. Since rats can only eat when water is available, there was a one-hour period immediately following testing during which the rat had full access to both food and water. The timeline of day to day testing is outlined in Figure 3B. When Sated, rats were allowed full access to both food and water for the full 24 hour period prior to testing. When Restricted of both food and water, the rat had access to both food and water for 1 hour after testing. On Food Restricted days, the rat had free access to both food and water for 1 hour, and then the food was removed, leaving only water for the 24-hour time period prior to testing. When the rat was Water Restricted, the rat had free access to food and water for one hour, and then the water was removed, leaving only food in the 24 hour period prior to testing.

Electrophysiology

Tetrode Implantation: To investigate neural firing patterns in multiple brain structures simultaneously, we used an electrode assembly consisting of 21 independently driveable "tetrodes." Each tetrode was made up of four 12.5μm Ni-Cr wires twisted together, which helps discriminate between individual neurons and gives a greater

capability for recording (Gage et al., 2010). Tetrode tips were gold-plated to lower impedances to 200-250 MΩ. Prior to implantation surgery, rats were taken off of food and water restriction for at least 24 hours. For drive implantation surgery, rats were anesthetized using isoflurane and a craniotomy was performed on the right hemisphere. Measurements were taken relative to bregma, so that individual tetrodes would reach the OFC, DMS, DLS and Nucleus Accumbens (both Core and Shell). Care was taken when removing the duramater to reduce brain swelling and bleeding. Skull screws were used to record electrocorticograms (ECoG) and served as a recording reference site, targeted on midline, approximately 1 mm posterior to lambda, and another ground, placed at the posterior lateral skull ridge. Additional skull screws were used to provide stability. After successful placement of the drive onto the brain surface, dental cement was used to fix the drive to the skull screws. As has been previously demonstrated (Berke et al., 2009; Berke et al., 2004), this drive set up allows us to record from a variety of brain structures simultaneously, and still allows the animal a free range of movement during recordings. In addition, each individual tetrode assembly was moveable, allowing for small adjustments within targeted structures to maximize neural recordings. Drive implantation surgery was performed on a total of 7 rats. Each 21-tetrode drive targeted the OFC, Insula, DMS, DLS, NAc Shell and NAc Core.

Recording Apparatus: The Berke lab utilizes custom-designed 96-channel systems, and records from all tetrodes simultaneously and continuously with minimal filtering. The drives weighed less than 15g, and did not impede behavior or movement during task performance. The cables that connected the drive to the computer were counterbalanced

and moveable, so the rat was free to move in any direction without restriction. During task performance, neural signals from each tetrode wire were recorded continuously at high-speed (31250 samples/s/channel), with minimal filtering ("wide-band", 1-9000Hz). The status of cue lights and photobeams was monitored at the same high speed (i.e. 32μs precision), and synchronized to the neural recordings.

Following a one-week recovery period from surgery, rats were tested daily (typically performing 150-400 trials in each ~2-hour session). Once recording began, a 4 day sequence of recordings was obtained before each tetrode was moved ~100μm to sample a new set of neurons for the next recording sequence, with at least 24 hours between tetrode moving and new recording to allow the brain tissue to settle.

Histology: At the end of the experiment, each tetrode site was marked with a small electrolytic lesion by passing a 25uA of current for 10 seconds, while under anesthesia. The rat was then perfused and the brain was Nissl stained so that the final tetrode locations could be mapped onto coordinates in a reference brain atlas (Paxinos and Watson, 2005), using Sqirlz Morph software (Xiberpix, Inc) and Matlab. From these atlas locations, prior recording day locations could be estimated by the number of screw turns. Example histological verification can be seen in Figure 3E.

Data Analysis

Spike sorting: The continuously digitized signal was first wavelet-filtered (Wiltschko et al., 2008) and then spike sorting was performed manually using Offline Sorter (Plexon INC, Dallas, TX). To differentiate unique cells, the difference in

waveform size and shape across all four tetrode wires was compared. Cells could be further classified as putative medium spiny neurons (MSN) or fast-spiking interneurons (FSI) based on clusters found in the scatter plot of two measurements of the wide-band spike waveform: the peak width at one-half maximum (FSI: 50-200ms; MSN 150-450ms) and the time from peak to valley (FSI: 100-455ms; MSN: 560-1500ms).

Behavioral Analysis: Behavior analysis examined free choice and forced choice trials separately. Procedural errors were excluded, and a two-tailed t-test was performed to test if choices on both free and forced choice trials were significantly different on Food and Water trials ($p< 0.05$). Additionally, trials were collapsed across days to examine the Preferred and Non-preferred behavior, so that Food trials on Food restricted days were Preferred, and Water trials on Water restricted days were Preferred. A similar behavioral analysis was then performed.

The median Initiation Times, Movement Times and Reward Retrieval Times were taken for each rat, and averaged across each trial type where the rat made it to the correct reward port (procedural errors and reward port errors were excluded). Analyses were performed separately on Food restricted and Water restricted days, as well as Sated and Restricted days. A two-tailed t-test was performed to test for significantly different reaction times between food and water trials ($p<0.05$).

Identification of trial related activity: After spike sorting and clustering, the population of cells was tested for trial-related activity. To reach criterion, a cell had to reach a minimum firing rate of 2 Hz on at least 2 trials of the same type (food-contra/food-ipsi,

etc.), within the same 30ms time window. An entire session was only included if the rat performed at least 80 trials, and was successful at >70% of the reward port entries. Analyses are limited to the days in which the rat was either food restricted or water restricted (unless otherwise noted).

Statistical Analysis: All analyses were performed using MATLAB (Mathworks, Inc, Natick, MA). To examine the population response, we performed a multiple regression analysis on each individual cell using the regstats function and the following formula:

$$\mathbf{S(t)= \beta_0 + \beta_1 A(t) + \beta_2 C(t) + \beta_3 T(t) + \beta_4 AX(t) + \beta_5 CX(t) + \beta_6 TX(t) + \varepsilon(t)}$$

where spike discharge rate S(t) for each trial (t) was analyzed using three task-related factors- Action (A(t)) (contralateral vs. ipsilateral), Goal port Chosen (C(t)) (food port vs. water port entered) and Cued Tone (T(t)) (food tone played vs. water tone played). $\varepsilon(t)$ indicates the error terms and β_0 β_6 are the regression coefficients. An additional regression analysis was run using the same formula, but instead had the factors Previous Outcome, Goal Chosen and Cued Tone

Each of the 6 behavioral events was analyzed separately, by centering a 6 second time window at the zero point of each event. A sliding window average of 500ms was used, moved in step sizes of 100ms for most analysis (unless otherwise stated). To better understand the specificity of encoding using the 3-factor regression, all combinations of main effects and interactions were examined from the 3-factor regression. From a first pass analysis, the combinations of main effects and interaction terms that best captured

encoding at reward port in were ‘Goal Only, ‘Current Outcome Only’, and ‘Goal+Cue+Current Outcome’. This last factor includes all cells that showed a main effect for Goal as well as a main effect for Cue, and had a significant interaction effect between those two terms. For action related firing, the ‘Action Only’ term captured most directional cells, while cells that encoded a main effect of ‘Action’, ‘Goal’ and the Interaction effect between the two also showed significant encoding. A diagram of the analysis used can be found in the Regression Triangle in Figure 3F.

To plot the proportion of cells from each brain region that reached significance in the 3-factor regression, an individual cell had to have a p-value < 0.05 for the main effect(s) and/or interaction effect(s) of the specific regression coefficient in a given 500ms window. The proportion of cells that reached significance was plotted across time for each separate combination of regression factors, centered on each of the 6 behavioral events. Finally, a binomial test was done with $p<0.05$ to test if the population encoding reached significance. A solid dot indicates the beginning of the 500ms time window that reached significance. The solid bars plotted across the top of the graphs indicate the start of a 500ms time window where at least 2 adjacent time-windows for a given region reach significance in the binomial test.

To further analyze what individual cells were firing for, we created a trial-type firing rate index. For a given regression factor, there were a set number of trial types that could potentially have the highest firing rate. For example, if a cell reached significance for ‘Goal Only’, it did so by differentiating between choosing the preferred vs. the non-preferred reward ports. If the average z-score firing rate was highest for a cell during a preferred-port chosen trial, then that cell was included in a total proportion of cells with

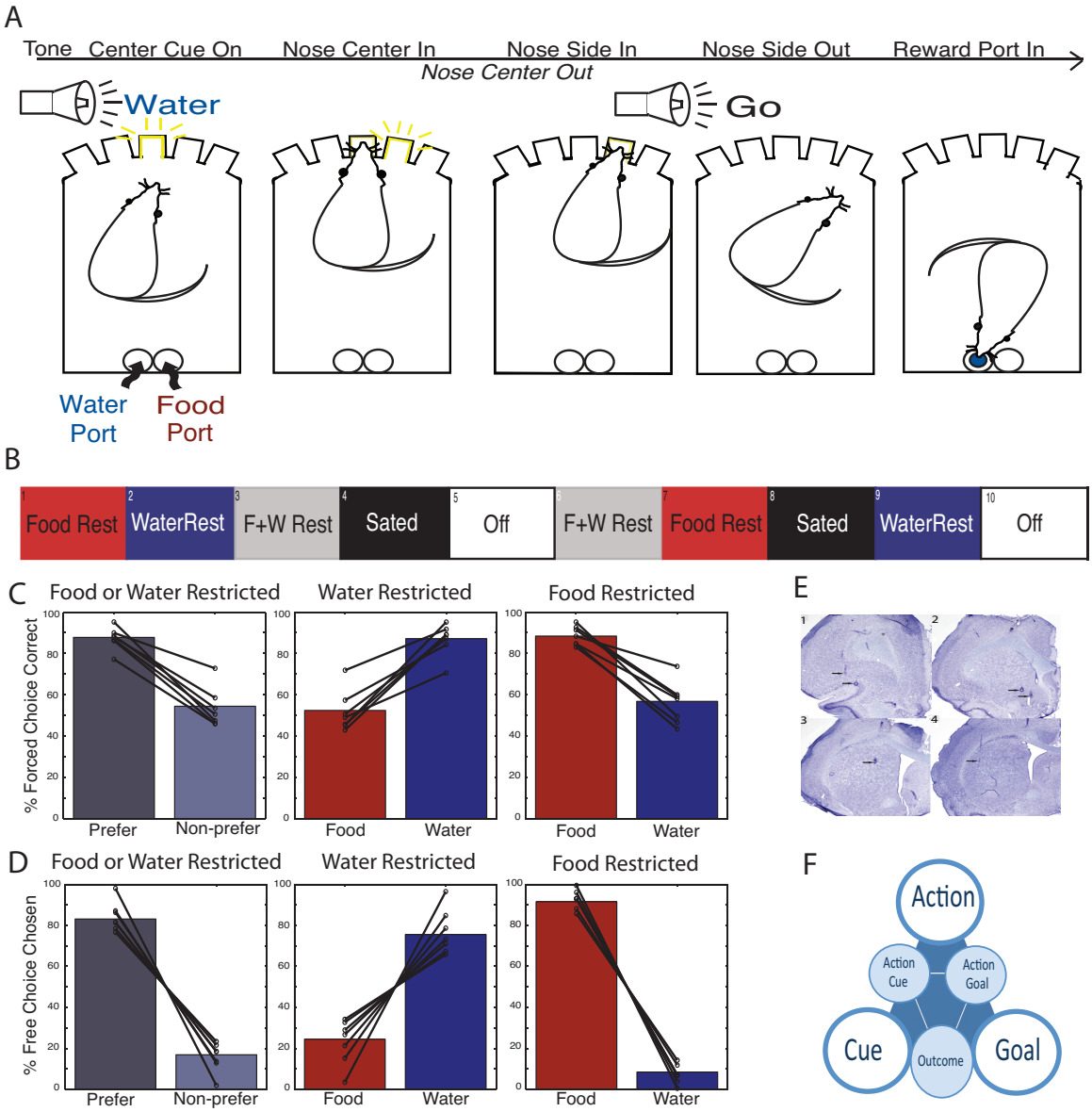
preferred-port chosen as the highest firing rate. The number of cells that had a higher firing rate for the non-preferred port was then subtracted from the number of preferred port cells to create the Value index. This same analysis was also done for Identity, and used water – food, rather than preferred vs. non-preferred, creating the Identity Index. A similar index was calculated for Action (contralateral – ipsilateral) and Outcome (correct – incorrect). Additionally, for cells that had 2 main effects, two indices are used- one for Value and another for Outcome. A positive number on the Value index means that there were more significant cells that had a higher firing rate for the preferred reward than the non-preferred reward. A negative number on the Outcome index means that there were more significant cells that had a higher firing rate for incorrect trials than correct trials (regardless of preference). To further demonstrate this Firing Rate Index, specific time points or an area under the curve was plotted as a bar graph, using specific time points related to the analysis.

To compare how each individual cell was encoding each of the different factors across time and across events, each cell from each region was individually plotted for all of the relevant regression factors. A 6 second time window was used around each event, with non-overlapping 200ms time bins. Each time a cell reached significance for a specific regression factor, that time bin was plotted in the color corresponding to the relevant factor. The Action/Goal/Cue plots have the cells sorted by the factor Goal, at the time of reward port in, and maintain that sorting across all of the behavioral events. The cell on line 1 stays the same for all behavioral events. The PreviousOutcome/Goal/Cue plot is sorted by the factor Previous Outcome at the time of the tone, and maintains that sorting across all behavioral events. To show the relationship between factors for the

entire region, the 3-factor regression analysis with the relevant factors is plotted across all behavioral events, in a similar manner described for Figures 7-14, but instead, each region is plotted separately (Figures 15-24). This was done for both 3-factor regressions.

Finally, individual cell overlap between events was examined. If a cell reached significance for 1 non-overlapping 200ms time bin used in the previous analysis, it was included. The specific factors examined were Action at 1 second surrounding nose side in, Goal at 1 second surrounding reward port in, Outcome at 0.5-1.5 seconds after reward port in, Integration of Cue+Goal from 1.5-2.5 seconds after reward port in, and Previous Outcome, -2 to -1 seconds before the tone. Cells that reached significance for the subsequent events and factors, as well as their overlaps are plotted using a Venn diagram. The outside circles represent cells that only reach significance in the corresponding factor/event, while the interaction circles are cells that overlap between factors and events.

Figure 3: Internal State Task and Behavior



A: Behavioral task outline. Each trial begins with a tone that indicates what reward the rat is working for (food, water or free choice). The rat pokes his nose in the lit center nose poke and then moves to an adjacent (lit) nose poke before moving to the opposite wall to retrieve a reward. The food and water ports are physically separated so the rat has to enter the correct reward port prior to receiving reward on forced choice trials, but can enter either port to receive the reward on free choice trials.

B: Motivational Manipulations. An example sequence of the motivational manipulations used to test internal state. This pattern was randomized for each rat.

C: Forced Choice Behavior follows internal state. On Forced Choice trials, the rats are instructed which reward to work for, and the rat only receives reward if he enters the correct goal port. Preferred and non-preferred refers to the relative value of the reward based on internal state (food when hungry, water when thirsty), shown as a combination of food and water restricted days together. Behavior is also shown broken down on Water

Restricted and Food Restricted days. Individual rat behavior is plotted as the solid black lines.
D: Free Choice Behavior follows internal state. Similar to figure (E) except that these show the choices the rat made when allowed to choose either food or water port on Free Choice Trials. A separate tone played indicating a Free Choice trial, and the rat received the reward from whichever goal port he entered first- there was no incorrect choice.
E: Histology Nissl staining sample images. Example tetrode placements in the OFC (1), NAC (2), DMS (3) and DLS (4). Full histology figures are in figure 6.
F: Regression Triangle, All Factors: Task-related Neural Integration. Each of the 3 main factors from the 3-factor regression analysis are shown with the interaction terms on the Regression Triangle.

Chapter 3: Results

Section 1: Behavior Results: *Internal motivation drives behavior*

Seven rats performed an average of 304 trials per day on food-restricted and water-restricted days, with a range of 140-488 trials per session. All seven rats used their internal state to guide decision-making on both free choice and forced choice trials, on food and water restricted days. When hungry (food restricted day), the preferred reward was food, and when thirsty (water restricted day), the preferred reward was water. On instructed, forced choice trials (75% of total trials), the rats correctly chose the preferred reward 86% of the time, while only correctly choosing the non-preferred reward 54% of the time, reflecting the fact that even when cued to go to the non-preferred port, the rats instead (incorrectly) chose the preferred port (Figure 3C, $p < 0.05$). On free choice trials (25% of total trials), when the rats could choose either reward port, the animals chose the preferred reward on 81.6% of trials and the non-preferred reward 18.3% of the time Figure 3D, $p < 0.05$).

The rats were allowed unlimited time to complete the nose pokes and get the reward, making it a self-paced task. Several time points in the task are behaviorally

relevant. These include the time from reward onset to first nosepoke, (trial initiation time), the time from first nosepoke to second nosepoke (movement time), and timefinal nosepoke exit to reward port entry (reward retrieval time) (Figure 3A, behavioral diagram).

Behavioral Metrics: Correct trials on either food restricted or water restricted days were used to examine how the rat moved through the task. The first relevant event was the trial initiation time- the time it took from when the rat heard the tone, to then poke his nose in the lit center port. In general, the rats were faster at this on Water Restricted days compared to Food Restricted days (Figure 4A and 4B). Examining just the Food Restricted days, it was surprising to find that the rats were actually significantly faster on Water trials, the non-preferred trial type, at initiation time (Figure 4B 3.9s vs. 3.7s). During the Movement Time and Reward Retrieval Time the rat moved faster on the food reward (the preferred reward). The reason that the Initiation Time for Water on Food Restricted days was faster than that for Food is because this analysis looked at only the correct trials. Examining all trials for reaction times shows that on all Water Cued trials, the rat initiates trials much more slowly, if incorrect trials are included (Figure 4C and 4D).

On Water Restricted days, the Initiation Times and Movement Times were both slightly, but not significantly faster on the Water Trials (preferred reward). The rats did move significantly faster during the Reward Retrieval Time on water trials, even though the difference in timing was very small (Figure 4B). Free Choice reaction times show a similar pattern to the Forced Choice reaction times (Figure 4C and 4D). In general, the

animals moved faster on Water Restricted days, for both reward types. Within Food or Water restricted days, the rats usually moved faster for the more preferred reward. However, this difference was not always significant, because the rat often moved just as quickly for a food reward on water restricted days (Figure 4D), which demonstrates that the rat may have been in a more highly motivated state when Water Restricted.

Examining Sated and Restricted for overall behavior shows that the rats performed slightly better on Restricted days for water trials, and slightly better on Sated days on Food trials (Figure 5A, n.s.). Free Choice behavior shows that on both Restricted and Sated days, Food was the more preferred reward (Figure 5B, $p<0.05$ for Sated days only). Individual rat preferences varied more on Free Choice trials on Restricted days, with some rats choosing Food more, and some rats choosing Water more (black bars on Figure 5B).

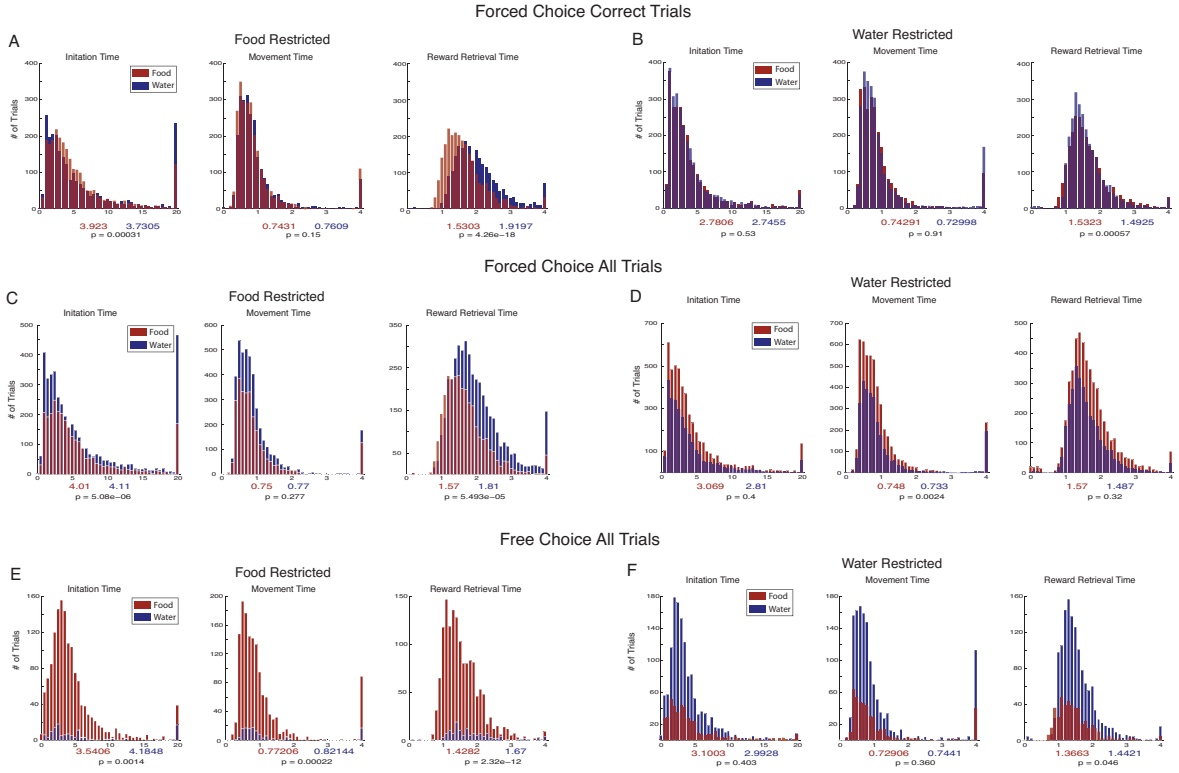
Reaction times on Sated and Restricted days for Forced Choice, correct trials shows the animals moved faster for food trials during the initiation and initial movement, but then faster on correct water trials when retrieving the reward, although these differences did not reach significance (Figure 5C). On Free Choice trials, the rats chose food or water with the same speed on each trial type, and moved at a similar speed as the Force choice trials (Figure 5D).

The behavioral choices and reaction times of the rat demonstrate that internal state influences how the rats perform. The biggest differences in the reaction time data are when the rat moves to receive the reward, with faster movements for the more preferred reward port.

Recording

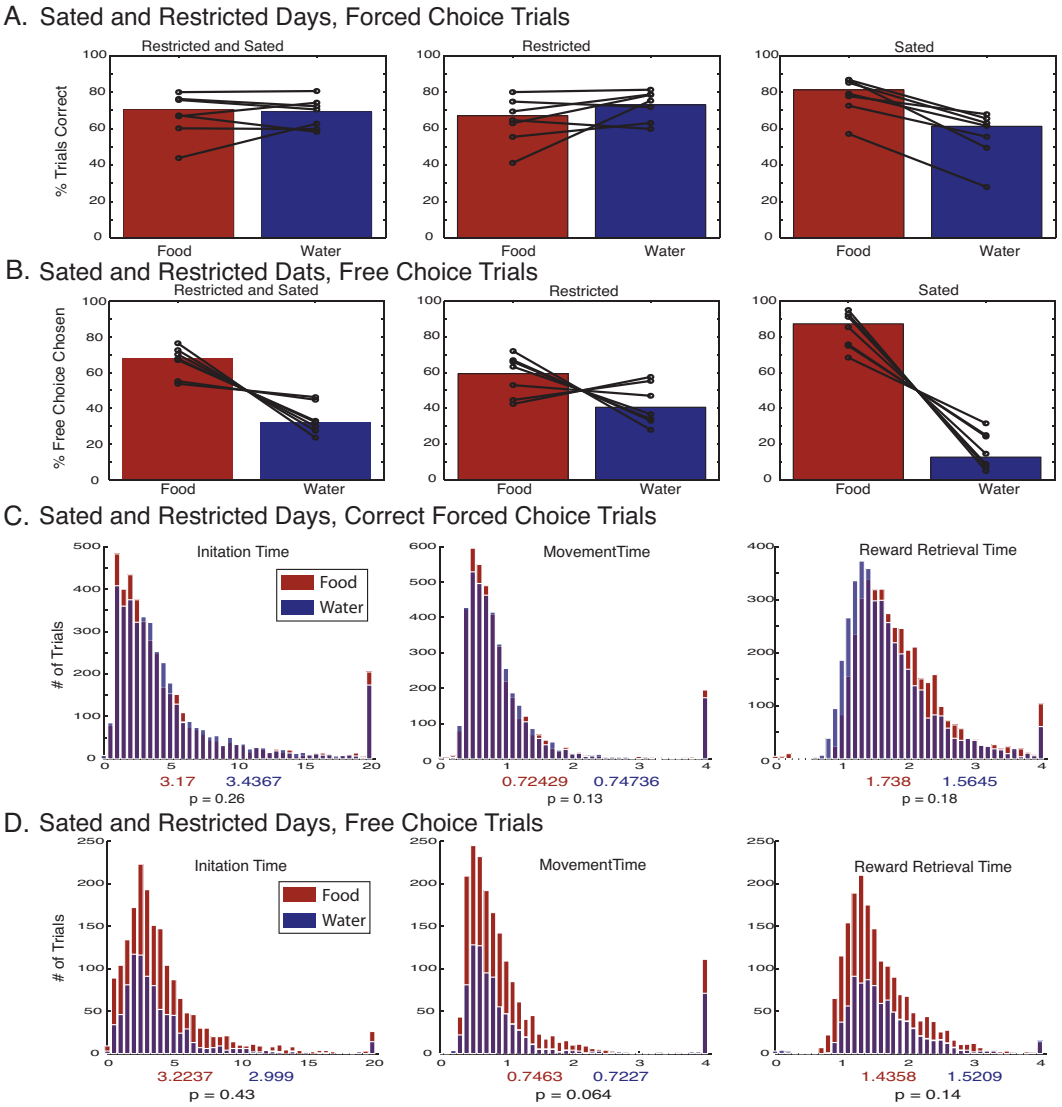
Single unit activity was recorded from a total of 2,345 cells over 141 days, which included of 605 cells from the DMS, 460 from the DLS, 335 from the OFC, 275 from the NAc Core, 572 from the NAc Shell and 98 from the Insula. In order to examine task-responsive neurons, we applied a firing rate criterion to each cell (see methods), and ended up with 1,103 neurons across 134 days of recording, with 127 cells from the DMS, 137 from the DLS, 115 from the OFC, 77 from the Core, and 101 from the Shell (Table 1). Insula recordings were not used for the remainder of the results. Histological verification of all of the recording locations for all of the cells used in the subsequent analysis can be seen in Figure 6.

Figure 4: Behavioral Metrics



A: Behavioral times on food restricted days for Forced Choice trials, using only correct trials. Initiation time is the time from when the tone plays until the rat pokes his nose in the first center nose poke. Movement time refers to the time it takes the rat to move from the center nose poke to the side nose poke. Reward retrieval time is the time it takes from when the rat leaves the last lit nose poke to enter the reward port. Food and water trials are plotted for all of the rats. Only correct trials were included in the reaction time analysis. Median times for food trials are listed in red, while median times for water trials are listed in blue. P-value for food vs. water trials are listed below the median times (t-test). Most behavioral times follow internal state, where times are faster for the preferred reward. Since correct trials are shown, the Cue and Choice are the same.
B: Behavioral times on water restricted days. Behavior and analysis is the same as in A.
C: Behavioral times on Food Restricted days for all trials, including incorrect trials. Analysis is the same as figures A and B except that behavioral metrics are shown for the trial that was Cued, regardless of if the rat was correct or not.
D: Behavioral times on Water Restricted days for all trials, including incorrect trials. The analysis is the same as in C.
E: Free Choice on Food Restricted days. The same analysis is used as in A, except that only Free Choice trials were analyzed.
F: Free Choice on Water Restricted days. The same analysis is used as in A, except that only Free Choice trials were analyzed.

Figure 5: Behavioral Metrics for Sated and Restricted days



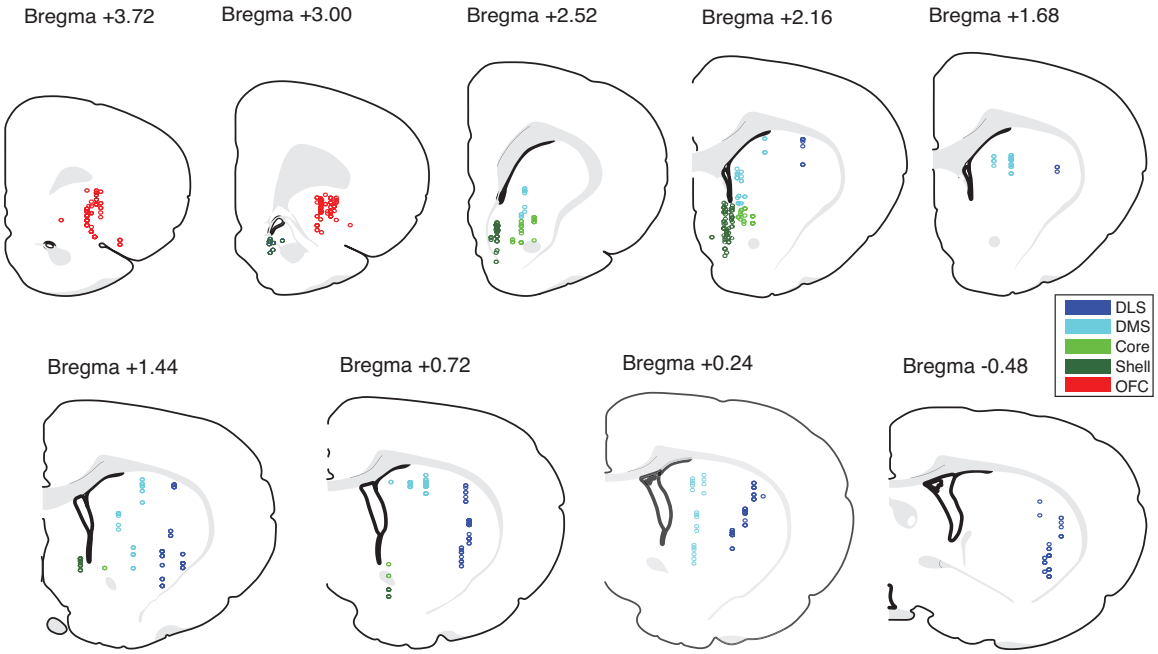
A: Forced choice trial behavior on sated and restricted days. Individual rats are shown in black lines.

B: Free choice trial behavior on sated and restricted days. Individual rats are shown in black lines.

C: All behavioral times for Forced choice trials, using only the correct trials on sated and restricted days. Same behavior and analysis as Figure 4.

D: All behavioral times on Free choice trials on Sated and Restricted days. Same behavior and analysis as Figure 4.

Figure 6: Location of all cells



The recording locations from each day, from each rat are plotted on the relevant brain atlases, listed by their AP coordinate taken from Bregma. Colored dots correspond to the brain region schematic used throughout the figures.

Table 1: Number of cells recorded from each rat, in each brain region

	DMS		DLS		OFC		NAC Core		NAC Shell		Totals
IM167	9	6	2	6	2	0	6	14	0	1	**46**
IM224	2	5	1	5	0	4	3	3	6	8	**37**
IM228	15	15	0	0	9	7	12	10	11	14	**93**
IM252	5	2	9	5	16	4	1	2	10	2	**56**
IM253	8	6	9	4	22	18	7	4	4	6	**88**
IM261	21	22	26	25	9	10	3	4	14	8	**142**
IM264	7	4	21	24	8	6	4	4	9	8	**95**
Totals	**67**	**60**	**68**	**69**	**66**	**49**	**36**	**41**	**54**	**47**	**557**
	127		**137**		**115**		**77**		**101**		

Section 2: Action Encoding: *Neural signal for action is distributed throughout the cortico-basal ganglia network*

This task was designed to disentangle specific aspects of reward-related behavior- the Cue indicating what reward is available, the movements taken to get to the reward, the reward port the rat chose, and the actual outcome of the trial. The main event corresponding to direction is 'nose center out', when the rat removes his nose from the center port, sees the lit adjacent port, and begins moving towards that port. In order to test our hypothesis that the dorsal striatum would have the highest proportion of cells encoding the direction of movement, we used a 3-factor multiple regression analysis with the three most relevant factors in the task: the cue-tone that played ('Cue'), the direction the rat moved ('Action') and the reward port the rat chose ('Goal') (See Regression Triangle, Figure 3F). For the Action factor, we were comparing movements made to the adjacent contralateral or ipsilateral nose poke port, and examined cells that only showed a significant main effect for Action.

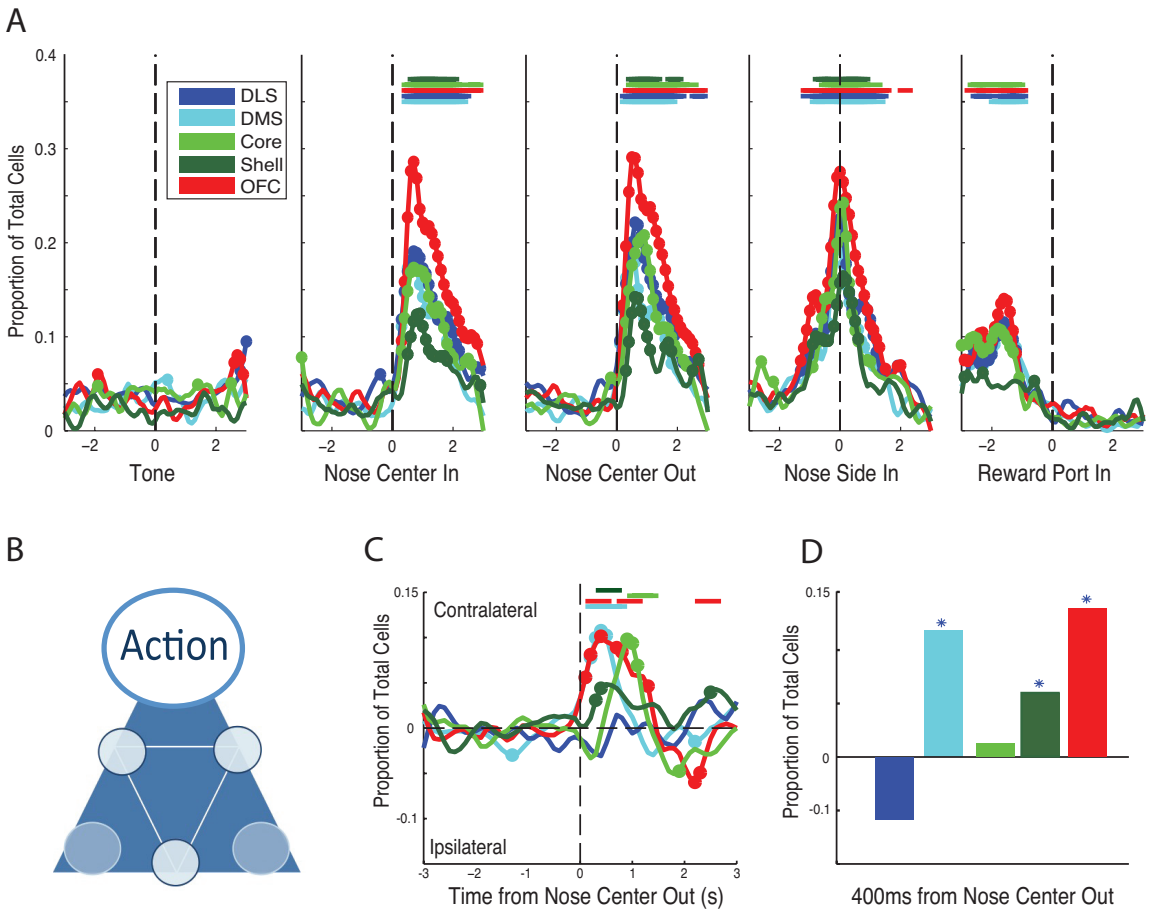
The original hypothesis was that only specific regions of the dorsal striatum- mainly the DLS- would encode a general motor response, due to the anatomical connections between the motor cortex and DLS (Alexander et al., 1986; Alexander et al., 1990; Voorn et al., 2004), and our single unit studies finding contralateral and ipsilateral

specific movements in the DLS (Gage et al., 2010; Schmidt et al., 2013). Previous rodent literature has found spatial encoding in the OFC, which may have been related to action (Feierstein et al., 2006; Furuyashiki et al., 2008), but no action-specific encoding has been found in the primate OFC (Padoa-Schioppa and Assad, 2006; Wallis and Miller, 2003), and no task has been utilized to specifically test action-related encoding in the OFC. Because of these findings, we expected to find very little encoding of Action in the OFC. In the Nucleus accumbens, previous literature has always focused on the integration of specific movements with a reward (Ito and Doya, 2009; Nicola et al., 2004a; Taha and Fields, 2006; Day et al., 2011). The few studies that have disentangled action from value (Goldstein et al., 2012) have focused on how value was encoded in the NAC, not movement. Due to these previous findings, we expected to find the greatest difference in contra/ipsi firing to be in the DLS, with very little encoding of movement in the other subregions.

Contrary to the hypothesis, all 5 brain regions had a significant proportion of cells that differentiated between contralateral and ipsilateral movements (binomial test, $p<0.05$), demonstrating that Action encoding was universal across the cortico-basal ganglia circuit (Figure 7A). Even more surprising was the fact that the brain region with the most Action-only encoding was the OFC, which reached a maximum proportion of 34%, at 500ms after nose center out. The DMS had 21% of cells at 500ms after nose center out, the DLS had 27% at 600ms, the Core reached a maximum of 25% at 900ms, while the Shell reached 18% at 600ms (Figure 7A). All 5 brain regions showed a similar timecourse of Action encoding, with the DMS and OFC reaching their maximum encoding earliest, while the Core reached its peak last.

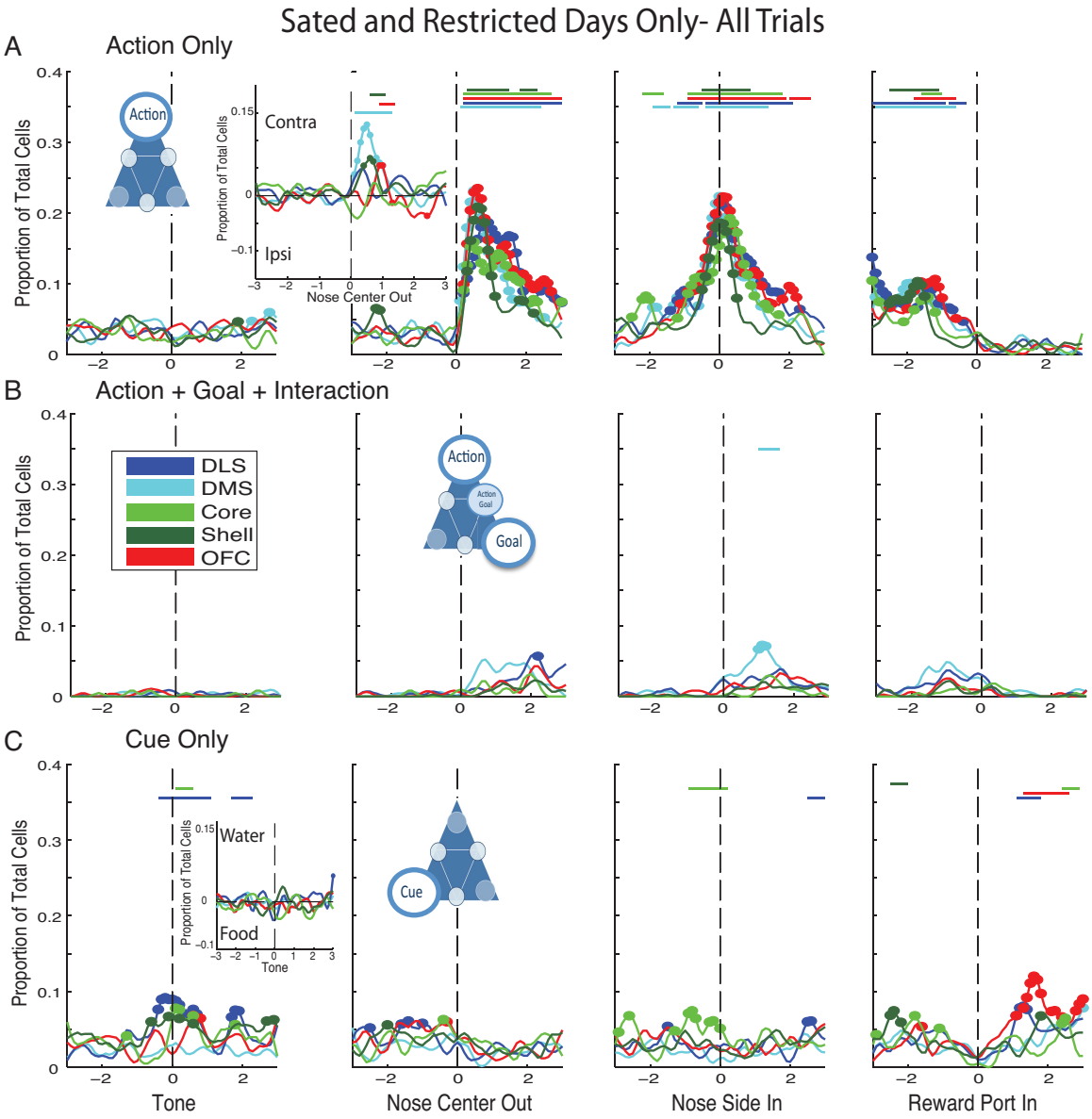
To test whether or not each region showed a bias in a specific movement direction (contra or ipsi), we examined the maximum firing rates in the two different trial types for each cell that reached significance in Figure 7A. Beginning as soon as the rat started moving, the OFC and DMS show significantly more cells with a higher firing rate for the contralateral movement (Figure 7C). The DMS had 20 cells (16%) firing more for the contralateral reward and 5 cells (4%) firing more for the ipsilateral reward (binomial test, $p<0.05$), and the OFC had 24 cells (21%) firing for contra vs. 8 cells (7%) firing for the ipsilateral side ($p<0.05$ binomial test). The Shell and Core also reach significance for a contralateral bias, but at a later time point (maximum bias in the Core at 900ms, and Shell at 600ms after nose center out). Surprisingly, the DLS does not show a significant directional bias, indicating that equal numbers of cells are firing more for both contralateral and ipsilateral movements. At 2.2 seconds after the rat begins moving, the OFC significantly changes its directional bias to the ipsilateral trials. Figure 7D shows the results from 400ms after nose center out from Figure 7C.

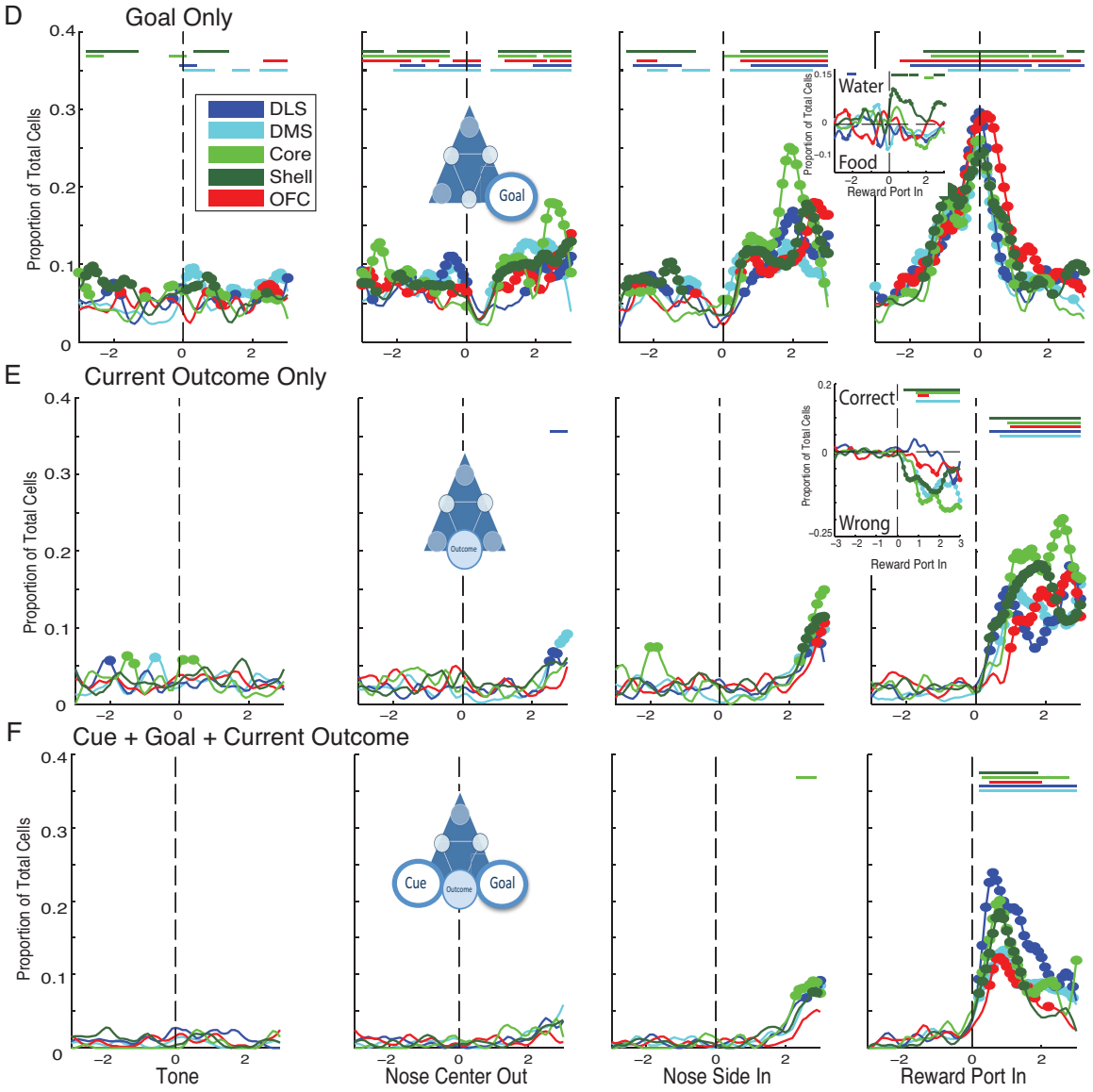
Figure 7: Action Encoding



A: Action Only Cells. Universal encoding of Action in all 5 brain regions. From the 3-factor regression (see formula, methods), the proportion of cells that reached significance for only the Action factor are plotted 6 seconds around each of the behavioral events, utilizing all trials (correct and incorrect). The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
B: Regression Triangle, Factor: Action Only.
C: Action Firing Rate Index. From the cells that reached significance in A, the trial type that had the highest firing rate is plotted as an index across time. Proportions are subtracted from one another to create the index: Contralateral – Ipsilateral. The solid dots and represents when a region, as a whole, reaches significance for one index (binomial test, $p<0.05$), while the solid bars across the top indicated when 2 adjacent bins reach significance. This shows a contralateral bias in DMS, OFC, Core and Shell, with no directional bias in the DLS.
D: Bar graph of the Action Firing Rate Index taken at 400ms from nose center out.

Figure 8: Sated and Restricted Days Only





A-F: The 3-factor regression analysis for factors Action, Goal and Cue are plotted on days when the rat was either sated for both rewards, or equally restricted for both rewards. The analysis is similar to that done in figures 7-10, with similar Identity Indexes also plotted for the firing rates. The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

A: Action Only

B: Cue Only

C: Action+Goal+Interaction

D: Goal Only

E: Current Outcome Only

F: Goal + Cue + Current Outcome

Section 3: Cue Encoding: *Neural signals for cue in the DLS and Nucleus Accumbens Shell*

Previous results have shown abundant cue-related activity in the NAc (Stalnaker et al., 2010; Roesch et al., 2009; Nicola et al., 2004b; Setlow et al., 2003; Taha et al., 2007; Ito and Doya, 2009). Therefore, the hypothesis was that there would be distinct groups of cells in the Nucleus accumbens Core, as well as the Shell, that would encode information related to the tone that played.

Lesion work in the Dorsal Striatum in rodents has led to the hypothesis that the lateral division of the dorsal striatum may be separately involved in stimulus-response behavior (Yin et al., 2004; Yin et al., 2006; Balleine et al., 2009). However, previous single unit studies from our lab and others have not consistently found a distinct stimulus-response, or 'Cue' encoding only in the DLS (Gage et al., 2010; Berke et al., 2009; Thorn et al., 2010; Stalnaker et al., 2010). Because of these conflicting results, we did not expect to find much cue specific encoding in the DLS, and none in the DMS. This study sought to examine how *specific* cue related information was encoded, though, differentiating this task from others that only examined general cue related activity. This task used 2 distinct tones that signaled different outcomes, allowing an analysis for the food vs. water tone.

On instructed trials, the tone ('Cue') signals which reward is potentially available upon correct completion of the current trial. The Cue indicates the same reward throughout both training and testing, and Tone presentation is independent of the rat's actions, or where the rat is in the recording chamber. The relative value of the Cue changes each day, depending on the rat's internal state. To test how the Cue is represented in the brain, the same 3-factor multiple regression described previously was used, but this time examined only cells that significantly differentiated between the two tones presented, on only the forced choice trials, regardless of the outcome of the trial. This 'Cue Only' factor reveals that both the DLS and Shell have a significant proportion of cells that encode information about the Cue, at the time of the tone (Figure 9A). At the time point of peak encoding for the factor 'Cue Only' (100ms), the DLS had 10% of cells encoding Cue, and the Shell had 11% cells encoding Cue. The OFC and DMS did not show a significant proportion of cells encoding a difference between the food tone and water tone, while the Core briefly reached an encoding of 12%, but only for one 500ms time window.

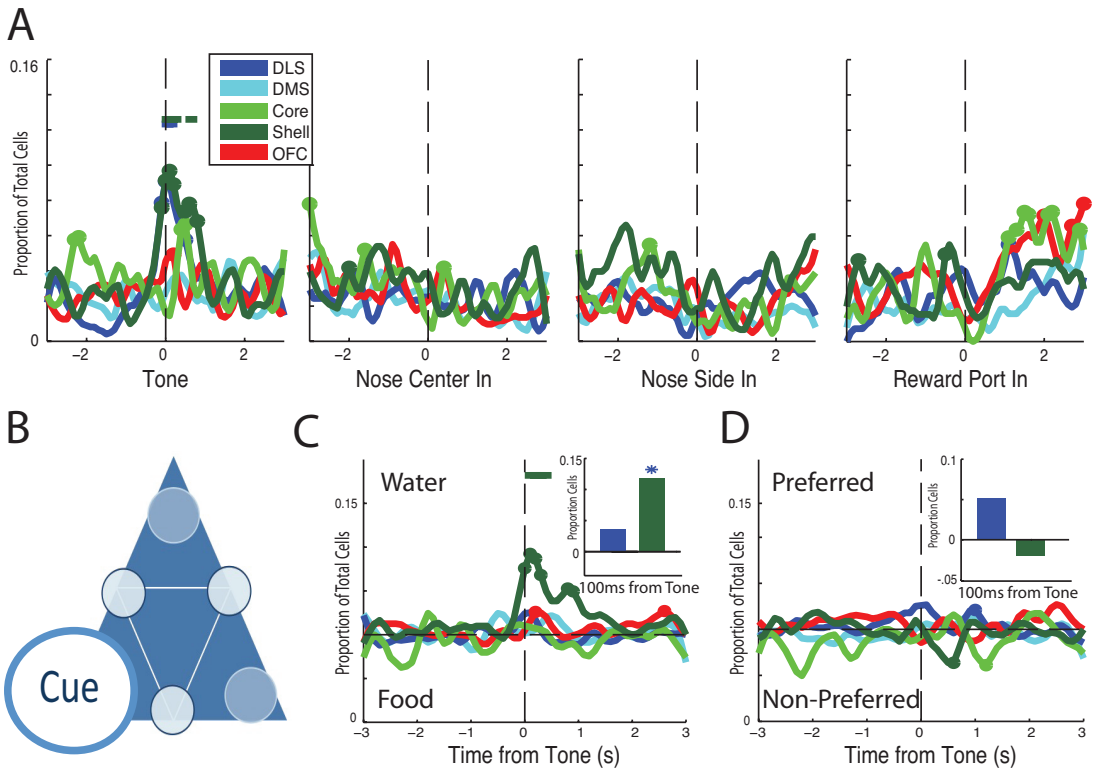
According to stimulus-response ideas of encoding, specifically in the DLS, it is theorized that the DLS is encoding this information without a representation of the goal in mind. This could also be viewed in a model-free framework, where the DLS is simply responding to a stimulus as it relates to the response, rather than a value/goal. If this were the case, we would expect to find equal proportions of cells firing for both food and water rewards, regardless of the current motivational state of the animal, meaning some cells would fire more for the water reward on BOTH food restricted and water restricted days. Previous findings in the NAC have typically found more firing for the more preferred

reward, so we hypothesized that any bias seen in the NAC subregions would be for water cue on thirsty days and food cue on hungry days.

The cells that reached significance for the factor Cue could be responding to the value of the potential reward, or they could be firing more for the reward it represents (food vs. water). To better understand how each region encoded Cue, a firing rate scale was used for Preferred vs. Non-preferred Tone (Value Scale) while a separate scale compared the tones in the context of Food vs. Water (Identity Scale). The difference in the two indices is that for the Value scale, the tones switch meanings based on whether the rat is hungry or thirsty, while the Identity scale remains the same regardless of internal state. Interestingly, the Shell tracked much better with the Identity scale, with all 11 cells (11%) firing more for the water tone than the food tone, ($p<0.05$, Figure 9C). By contrast, the DLS cells encoding Cue showed a trend towards the Value scale, with 10 cells (7%) firing more on preferred trials and 4 cells (3%) firing more for non-preferred trials 200ms after the tone (Figure 9D, $p=0.09$). The Core only showed a very brief bias for the Non-Preferred Cue, but this bias was not at the same time point when the region, as a whole, reached significance.

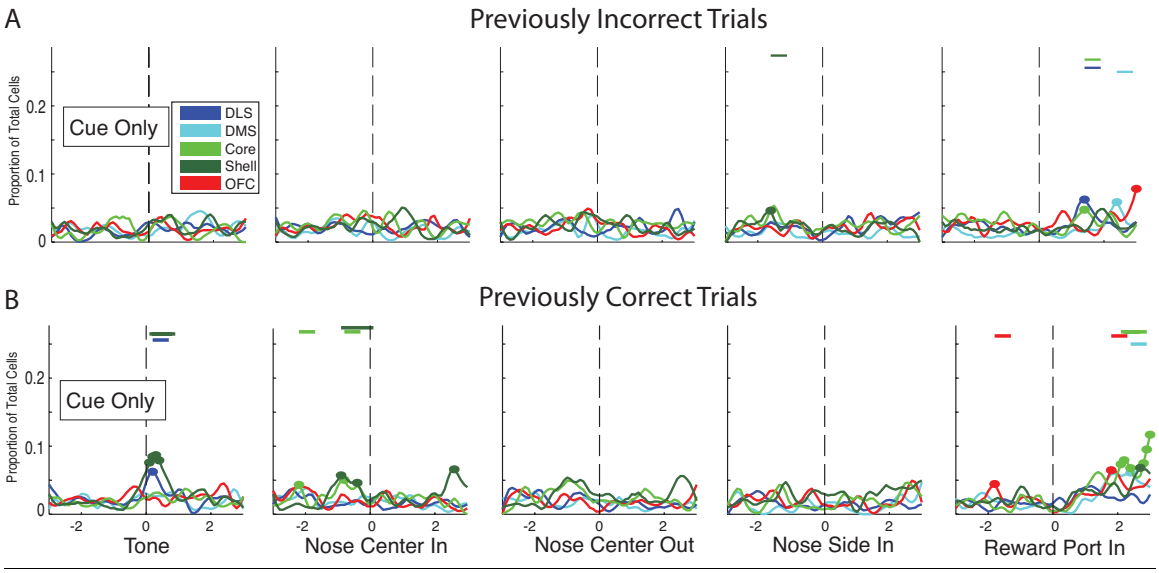
The results of finding no specific bias in the DLS confirmed the hypothesis that the DLS is most likely encoding the Cue in a model-free, stimulus-response based way. There was no specific increase in encoding for either a specific food/water tone, or a specific preferred/non-preferred tone- the DLS was simply encoding the stimulus. Contrary to our hypothesis, the Shell on the other hand, was only firing for the tone that signaled water.

Figure 9: Cue Encoding



A: Cue Only Cells. From the 3-factor regression, the proportion of cells that reached significance for only the Cue factor are plotted 6 seconds around each of the behavioral events, utilizing all trials (correct and incorrect).. The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 300ms, sliding in 100ms steps. Cue encoding only in DLS and Shell show striatal specialization.
B: Regression Triangle, Factor: Cue Only.
C: Cue Identity Firing Rate Index. From the cells in A, the trial type that had the highest firing rate is plotted as an index across time. Proportions are subtracted from one another to create the index: Water – Food. The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 300ms, sliding in 100ms steps. Inset shows the proportion of cells 100ms after the tone for Water (top) and Food (bottom, no cells). Firing rate bias is for Identity in the Shell.
D: Cue Value Firing Rate Index. From the cells in A, the trial type that had the highest firing rate is plotted as an index across time, but trials are identified by preference, so food on food restricted days and water on water restricted days. Proportions are subtracted from one another to create the index: Preferred – NonPreferred. The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Inset shows the proportion of cells 100ms after the tone for Preferred (top) and Non-preferred (bottom). There is no firing rate bias for Value in any brain region.

Figure 10: Cue Only on Previously Incorrect or Correct Trials



A: Cue for Previously Incorrect Trials: The 3-factor regression for Action, Goal and Cue is shown for the Cue only factor, using trials that had previously been incorrect. In this behavioral task, when a trial is incorrect, it is repeated on the next trial. The filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 300ms, sliding in 100ms steps.
B: Cue for Previously Correct Trials. The cells that reached significance for Cue Only in the 3-factor regression for Action, Goal and Cue is shown using trials that had previously been correct. Since in this behavioral task, an incorrect trial is repeated, this means that previously correct trials are unique, and have a new tone presented. Response to the Cue happens on 'unique' trials, when there is no information known to the rat, prior to the tone playing.

Section 4: Action-Goal Encoding: *Dorsal medial striatum drives action-goal encoding*

Since this task specifically disentangles action from the goal port, it can help answer the question- which brain region encodes both action and goal? The reward tone signals which reward port the rat must enter, but the directional cue is only signaled via lit nose ports as the rat is moving- meaning directional and reward cues are separate and unrelated. In addition, the goal port the rat actually chooses does not always correspond to the port that was signaled, leading to the more precise measurement of Goal. That is, the Goal the rat chose does not guarantee reward, so any signal related to the upcoming Goal is not simply an anticipation of reward delivery. To examine the interaction of Action and Goal, the same 3-factor regression was used with variables direction (contralateral vs. ipsilateral), Goal (food vs. water CHOSEN) and Cue (food vs. water TONE). All trials on food and water restricted days were used, regardless of if the rat chose the correct reward. From the 3-factor regression, only cells that showed a main effect for 'Action', a main effect for 'Goal', and an interaction between the two terms were analyzed (Figure 11A inset). Cells that only had an interaction effect, or just one main effect and the interaction effect were not included in this analysis, and no brain

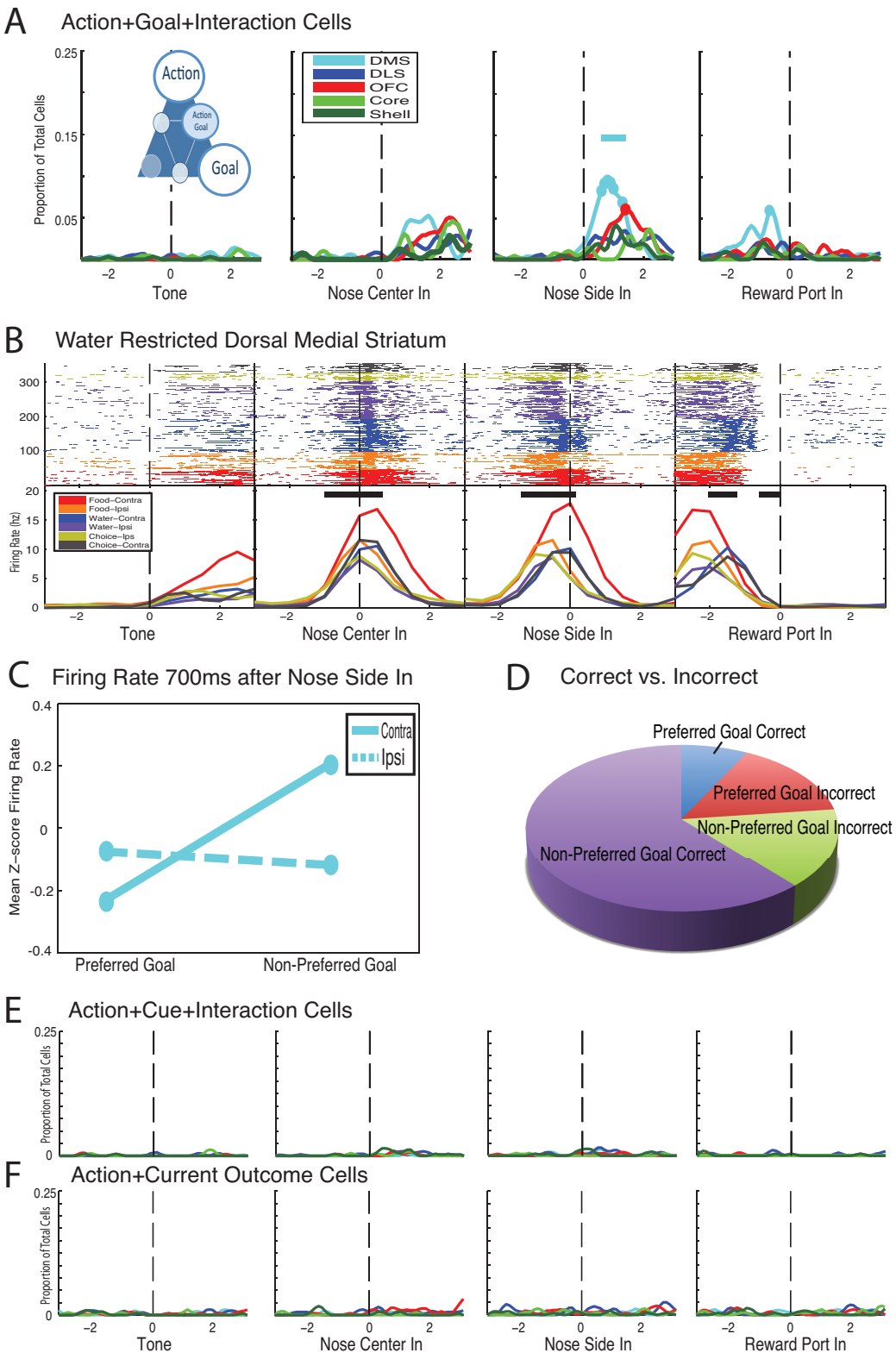
region reached significance for any combination of factors other than those in Figure 11A.

Previous work has found Action-Value like encoding in the striatum of primates and rodent (Kim et al., 2009; Kimchi and Laubach, 2009b; Samejima et al., 2005; Seo et al., 2012; Pasquereau et al., 2007; Lau and Glimcher, 2008), as well as some encoding in the NAC in rodents (Roesch et al., 2009). However, previous research has specifically focused on how the brain integrates action with value. The current task was designed to separate those two variables. If the brain still integrates action with a goal, or value, then it would be in cells that actually participate in both of those aspects of responding, separately, as well as becoming engaged in the task when those aspects are integrated. Rather than examining how the individual regions were working to bias a response for the more preferred reward, this analysis is more similar to an action-selection task, since there were no action + value integrations, and the reward likelihood was a known value to the rat.

As hypothesized, the strongest encoding of action-goal occurs in the DMS, where a maximum of 10% of cells reached significance for Action, Goal and the interaction term (Figure 11A), at 700ms after nose side in. An example cell from the DMS is shown in Figure 11B. To understand how the DMS was encoding Action-Goal, the average z-score firing rates of the significant cells in Figure 11A were plotted at the time window of maximum encoding (700ms after nose side in) in Figure 11C. This plot uses the firing rates from each of the 4 different trial types- contralateral-preferred, contralateral-non-preferred, ipsilateral-preferred, ipsilateral-non-preferred. Using this analysis, it is clear that the DMS is firing the most for contralateral + non-preferred goal.

The average firing rates from the DMS cells demonstrate that as a region, firing is enhanced on trials when the rat chooses the non-preferred reward and moves in the contralateral direction. An example DMS cell that is firing for the non-preferred reward (food) and the contralateral direction (left) is shown in Figure 11B. The DMS is the only region that reaches significance with this combination of factors. Taking the cells that were plotted in Figure 11B, we also examined if the firing rate was highest on correct or incorrect trials, and plotted those proportions in Figure 11D.

Figure 11: Action Goal Cells in the Dorsal Medial Striatum



A: Action + Goal + Interaction cells. From the 3-factor regression, the proportion of cells that reached significance for the main effect of Action, Goal and the interaction term between the two are plotted 6 seconds around each of the behavioral events, utilizing all trials (correct and incorrect). The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps. Striatal specialization is seen in the DMS for Action + Goal encoding.
B: An example cell from figure A is shown. This cell fires most for the food-contralateral trials on a water-restricted day (non-preferred). The top is the raster plot, and the bottom panel is the peri-event time histogram. The black bar represents when the cell reached significance for Action + Goal + Interaction.
C: The mean z-score firing rate from cells that reached significance at 700ms after Nose Side In from A, only in the DMS, are plotted based on the goal choice as preferred vs. non-preferred, on ipsilateral vs. contralateral actions. This shows that the firing rate in these cells was highest on contralateral movements to the non-preferred goal port.
D: The cells that were plotted in B are broken down by trial type, as correct vs. incorrect for the preferred vs. non-preferred reward. A majority of cells had the highest firing rate for the non-preferred correct goal port.
E: Action + Cue + Interaction (action-cue). From the 3-factor regression, this demonstrates that no brain regions reached significance for the main effect of Action, Cue and the interaction term between the two.
F: Action + Current Outcome (goal-cue). From the 3-factor regression, this demonstrates that no brain regions reached significance for the main effect of Action, while simultaneously reaching significance for the interaction term between the Cue and Goal (aka Outcome).

Section 5: Reward Port Encoding: *Universal goal encoding, with subregional specializations*

The final stage of the task occurs when the rat leaves the last nose poke port and enters the reward port: the 'Goal'. At this point the rat can still get the trial correct or incorrect because he still has to enter the correctly cued port first, before he receives the reward. Previous research has found upcoming goal choice in the OFC (van Duuren et al., 2007; Sul et al., 2010; McDannald et al., 2012), and the dorsal striatum (van der Meer et al., 2010). It is unclear whether or not impending goal choice is always encoded in the ventral striatum (van der Meer and Redish, 2010; Kim et al., 2007). We used the 3-factor regression and examined cells that were significant for Cue-only, Goal-only, and Cue-Goal Interaction. Surprisingly, we found that all 5 brain regions had a significant proportion of cells encoding Goal-only, beginning almost 2 seconds before entering the final port (Figure 12A), while no brain regions reached significance for the Cued Reward at the time of reward port in (Figure 9A), indicating that the cells were differentiating where the rat actually went versus where the rat was instructed to go.

Goal encoding was universal across regions, but the relative proportions and timing of maximum encoding differed between areas. Prior to entering the reward port, the highest proportion of coding cells was found in the OFC (18%) and Shell (17%), followed by DLS (15%), DMS (14%), and Core (12%). In contrast, *after* reward port in,

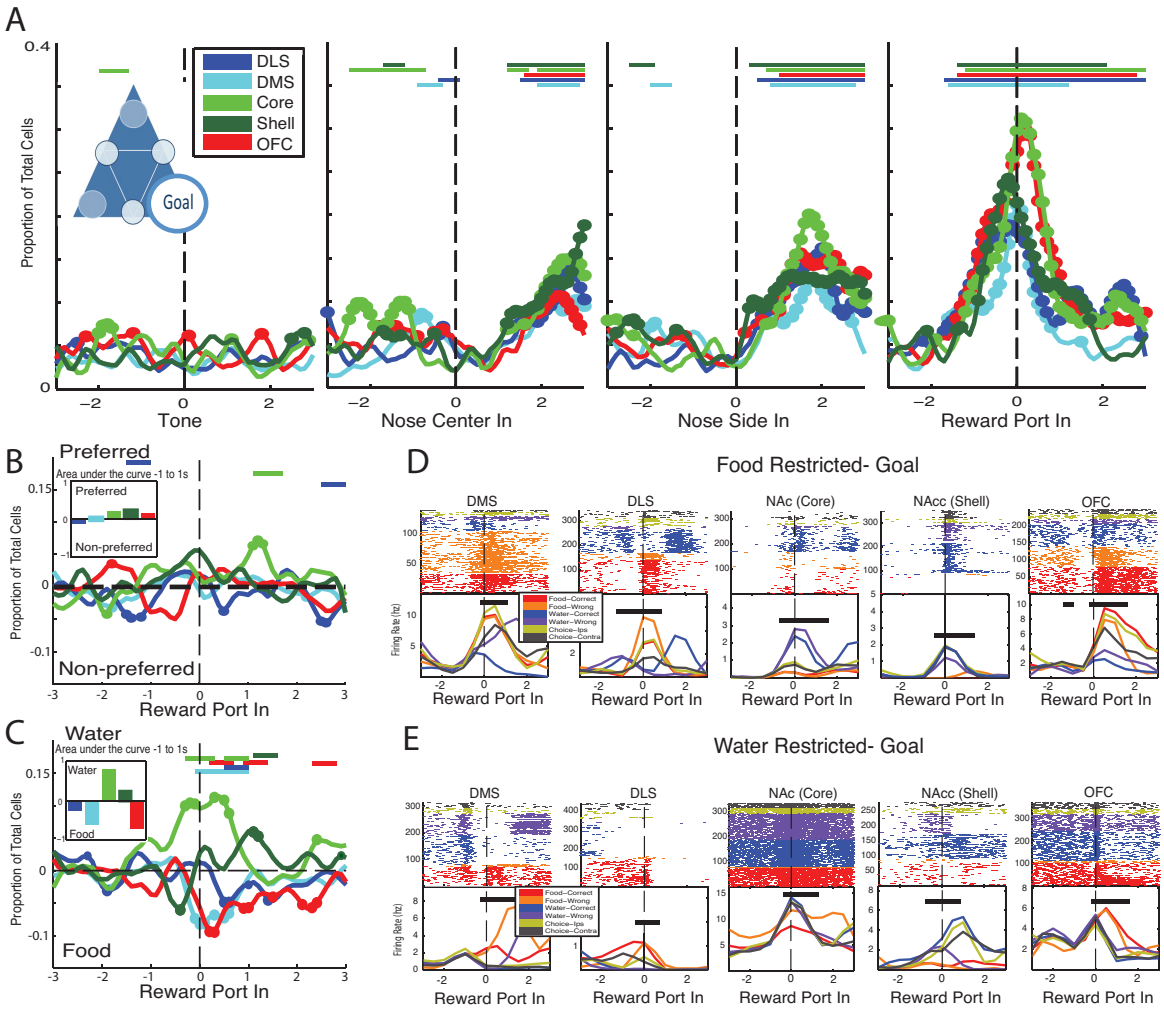
the Goal port was encoded in 27% of Core units, followed by OFC (18%), DLS (12%), DMS (11%), and Shell (11%). Peak encoding also demonstrates the different profiles between the 5 regions: the Shell and DLS peak *before* reward port in (26% at -300ms and 20% at -100ms), while the DMS and Core peak 100ms after reward port in (23% and 32 % respectively), and the OFC peaking latest, 300ms after reward port in (34%).

The cells encoding the factor Goal-Only show differentiation between which reward port the rat enters. It was our hypothesis that certain brain regions would show a bias for the higher value port (e.g. the food port when hungry), by having a higher firing rate before and after entering that port. To test this hypothesis, we used the Value Index for preferred vs. non-preferred trial type firing rates (see methods). Contrary to our hypothesis, there was no single brain region with a bias for the higher value port (Figure 12B). Next we used the Identity Index to test for encoding of the food port vs. the water port, regardless of internal state. Unexpectedly, we found regional biases for the identity of the port (Food vs. Water port, Figure 12C). The Core reaches significance before reward port entry for the Water port, while the DMS reaches significance before Food port entry. The OFC reaches significance for the food port after reward entry, and the Shell reaches significance for the water port long after reward entry. The bar graph insets in Figures 12B and 12C demonstrate the total encoding of the respective firing rate indices from 0 to 1 second after reward port entry, which again shows that there was no regional bias for value, but there were strong regional biases for the identity of the reward port. Despite previous research showing higher firing rates and regional biases for higher valued or more preferred goals, our results indicate that the identity of the reward port

was the only factor to show a bias in some brain regions, and only after reward port entry, despite the fact that all 5 brain regions had strong encoding of the upcoming goal.

Individual cell examples are shown in Figures 12D and 12E on either food or water restricted days. The DMS, DLS and OFC have 2 different cells that fire most when the rat enters the food port on either day. The Core and Shell show higher firing rates when entering the water port, regardless of internal state.

Figure 12: Goal Only Cells



A: Goal Only Cells. From the 3-factor regression, the proportion of cells that reached significance for only the Goal factor are plotted across behavioral events. The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps. The inset triangle shows the Goal factor on the Regression Triangle. Universal encoding for the port the rat enters is seen prior to entry in all 5 brain regions.

B: Value Firing Rate Index. From the cells in 9A, the trial type that had the highest firing rate is plotted as an index across time, but trials are identified by preference, so food on food restricted days and water on water restricted days (Preferred – NonPreferred). The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Inset shows the area under the curve of the line plotted in B, from -1 to 1s after reward port in. There is no firing rate bias in any brain region for the Value of the port that is entered.

C: Identity Firing Rate Index. Same as B except that each cells is plotted for Identity, so the index is Water – Food. The inset shows the area under the curve from -1 to 1s after reward port in. There is a bias in all 5 brain regions for the port the rat enters, based on goal identity.
D: Example cells from each of the 5 brain regions on a Food Restricted day. Black bar indicates when the cell reached significance for encoding the Goal factor. The top plots are the raster plots while the bottom plots are the peri-event time histograms
E: Example cells from each of the 5 brain regions on a Water Restricted day. Black bar indicates when the cell reached significance for encoding the Goal factor.

Section 6: ***Outcome Encoding After Reward Port In***

The interaction term between Goal and Cue represents the 'Outcome' variable, and indicates whether or not the rat was correct or incorrect. Each of the 5 brain regions had cells that showed significance for the interaction term only, without any of the 2 main effects, indicating that a signal about whether or not a trial was correct was universal. However, no region began significantly encoding the Outcome until after the rat entered the reward port (Figure 13A). Despite a widespread signal for Outcome, this signal was strongest in the Core, although it began at a later time point than most of the other regions (Figure 13A). The maximum proportion of cells encoding Outcome reached 26% in the Core, at 2.7 seconds after reward port in. The DMS reached 17% (+1.7 seconds), the DLS reached 16% (+1.9seconds), the OFC reached 16% (+2.2seconds) and the Shell reached 18% (+2.9seconds).

Outcome encoding showed the strongest regional bias on the firing rate index among all 5 brain regions, compared to all of the other factors examined. To quantify this bias, we looked at the firing rates on correct vs. incorrect trials and created the Outcome Firing Rate Index (Figure 13B). All 5 brain regions reached significance 700ms after entering the reward port. The DLS, as a region, had a significant bias to fire more on correct trials, just after reward port in (9 cells firing more for the correct reward and 2 cells firing more for the incorrect reward, Figure 13B $p<0.05$ binomial test). All of the other brain regions had a significant bias for the incorrect outcome- the DMS, OFC, Core and Shell each only had 1 cell fire more on correct trials (Figure 13B, $p<0.05$), with significantly more cells firing at a higher rate for the incorrect outcome. The DLS pattern

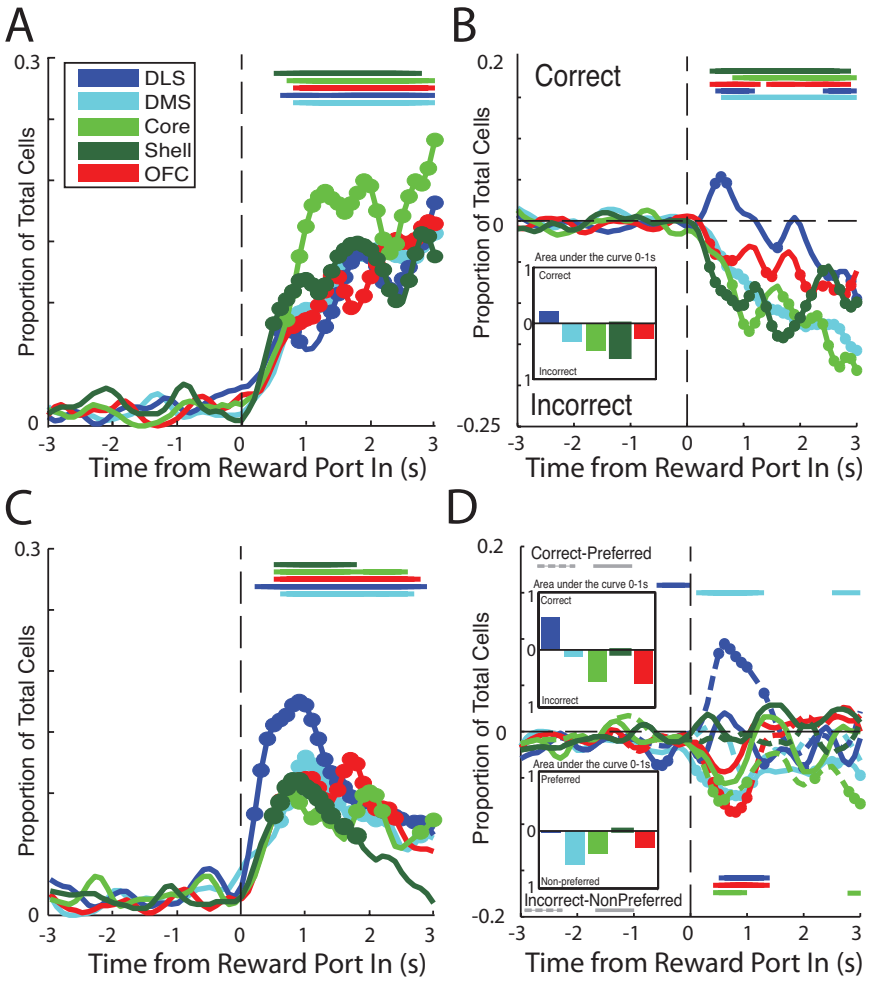
changed 2.2 seconds after reward, when the DLS then had 5 cells firing for the correct outcome, and 10 cells firing for the incorrect outcome (Figure 13B).

The findings for Outcome were nearly the same on sated/restricted days as they were on food or water restricted days, with the Core showing the highest proportion of cells encoding Outcome (Figure 8E). The DMS, OFC, Core and Shell all had significantly more bias for the incorrect reward, while the DLS had a trend towards the correct reward (Figure 8E inset).

The final combination of factors that showed significance at the time of the reward were cells that had a main effect of Cue, a main effect of Goal, and a significant interaction term between Cue and Goal. The DLS showed the greatest proportion of cells for this term, as well as encoding these variables earlier than all other brain regions. At 1 second after reward port in, the DLS reached a proportion of 18%, the DMS at 15%, the OFC at 10%, the Shell at 9% and the Core at 7% (Figure 13C).

The firing rate scale for this combination of terms again shows that the DLS is biased to fire for the correct reward, up until 1.5 seconds after reward port in, while the other 5 brain regions show a bias for the incorrect reward, though this only reached significance in the OFC and Core (Figure 13D, solid lines). Interestingly, the DMS again had a higher firing rate for the non-preferred goal (Figure 13D, dashed line) and was the only region to reach significance for the Value firing rate scale, demonstrating that the DMS was the only region with a bias for value, and did so by firing more for the non-preferred reward. This same analysis for Stimulus (food vs. water) did not yield a significant bias in any brain region.

Figure 13: Reward Integration Cells After Reward Port In



A: Current Outcome Only cells. From the 3-factor regression, these are cells that reached significance for the interaction between Goal and Cue only, plotted 6 seconds surrounding reward port in. The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps. Universal coding of the Outcome is seen in all 5 brain regions, with the Core reaching the highest proportion of cells.

B: Outcome Firing Rate Index. From the cells in A, the trial type that had the highest firing rate is plotted as an index across time. Trials are identified by Outcome, (Correct - Incorrect). The solid dots and corresponding bar represents when a region, as a whole, reaches significance for one index (binomial test, $p<0.05$). Inset shows the area under the curve of the line plotted in B, from 0 to 1s after reward port in. The DLS is the only region to show a bias for the correct outcome- all other brain regions are biased for the incorrect outcome.

C: Reward Integration Cells. From the 3-factor regression, these are cells that reached significance for the main effect of Goal, the main effect of Cue and the interaction between Goal and Cue, plotted 6 seconds surrounding reward port in. The filled circles

represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
D: Outcome and Value Firing Rate Index. From the cells in C, the trial type that had the highest firing rate is plotted as an index across time. Trials are identified by Outcome for the dashed lines, (Correct - Incorrect), and Value for the solid lines, represented as Preferred – Non-preferred. The solid dots represents when a region, as a whole, reaches significance for one index (binomial test, $p<0.05$). The solid bars on the top half of the plot correspond to significance on the Outcome Index, with a corresponding inset showing the area under the curve for this index from 0-1s after reward port in. The solid bars on the bottom half of the graph show significance for the Value index, with the inset showing the area under the curve of the line plotted from 0 to 1s after reward port in.

Section 7: *Reward History*

The rat's performance on the previous trial determined if the current trial was a repeat of the previous trial, or if the current trial presented a unique tone to the animal. Since trial history affected the presentation of each trial, we analyzed how this encoding was represented in the brain, as Reward History. This analysis followed the encoding of the current outcome factor into the inter-trial interval and into the next trial. This separate regression analysis examined the effect of the previous outcome on firing rates seen in the current trial. This is especially interesting because the task was designed so that after an incorrect trial, the rat had to repeat that trial. Therefore on some trials, the rat already knew what the trial would be before the cue even came on. To analyze how reward history was encoded in the brain, a 3-factor multiple regression analysis was performed using the factors 'Previous Outcome', 'Cue', and 'Goal'. The proportion of cells from each brain region that encoded only Previous Outcome are shown in Figure 14A across all behavioral events.

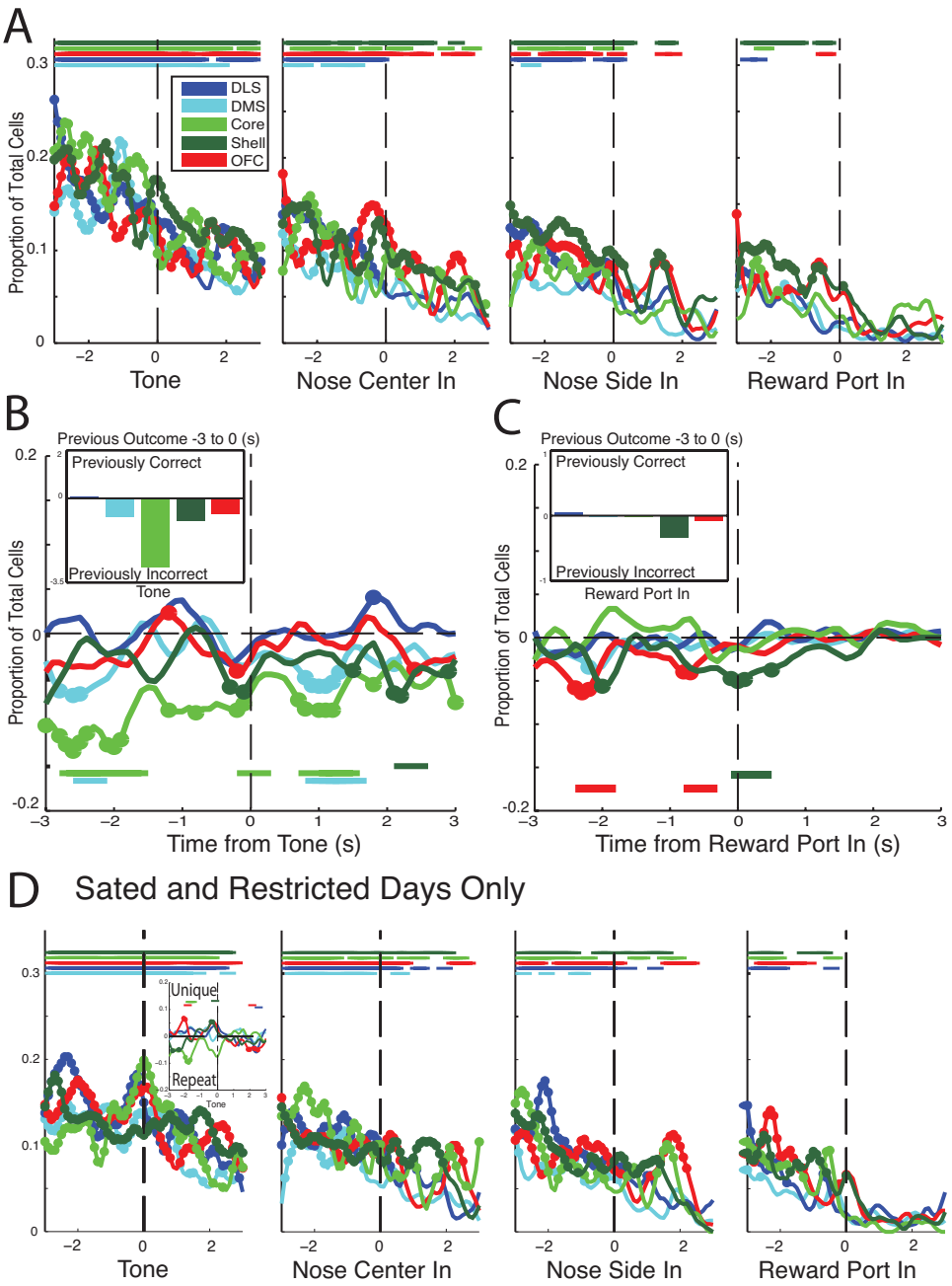
In order to determine how the individual cells were encoding the previous outcome, a firing rate index was plotted for Previous Outcome, and correct vs. incorrect trials were examined at two relevant time points, the Tone and at Reward Port In. In this analysis, the Core fires significantly more for the previously incorrect trials in the 3 seconds prior to the tone (Figure 14B). No other region has a significant bias prior to the tone coming on. As the rat approaches the reward port no regions show a significant bias for either previously correct or incorrect trials (Figure 14C).

To test if there were other aspects of the previous trial that influenced firing rate, other multiple regression analysis were run for Previous Outcome + Previous Goal,

Previous Goal + Current Outcome, and Previous Outcome + Current Outcome. None of these interactions reached significance for any of the brain regions.

The final analysis was examining Previous Outcome on sated and restricted days, when the value of the two rewards was similar. When we ran the 3-factor regression using Previous Outcome, the same pattern of response was seen in all 5 brain regions as on the food and water restricted days, with the Core showing the strongest effect of Previous Outcome. The Core also fired more for the previously incorrect reward prior to the Cue in the current trial (Figure 14D).

Figure 14: Reward History



A: Previous Outcome Only cells. From a new 3-factor regression examining Previous Outcome, Goal and Cue, these are cells that reached significance for the factor Previous Outcome Only, plotted 6 seconds around each of the behavioral events. The filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps. All 5 brain regions significantly encode whether or not the previous trial was rewarded, prior to the tone playing in the current trial.

B: Previous Outcome Firing Rate Index. From the cells in A, the trial type that had the highest firing rate is plotted as an index across time. Trials are identified by whether or not the previous outcome was correct (Previous Correct - Previous Incorrect). The filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Inset shows the area under the curve of the line plotted in B, from -3 to 1s before Tone.
C: Previous Outcome Firing Rate Index. From the cells in A, the trial type that had the highest firing rate is plotted as an index across time. Trials are identified by whether or not the previous outcome was correct (Previous Correct - Previous Incorrect). The filled circles represent when a region reached significance (binomial test, $p< 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Inset shows the area under the curve of the line plotted in B, from -3 to 1s before reward port in.
D: Previous Outcome plotted on Sated and Restricted days only. Same analysis used in 14A.

Section 8*: Cell and Event Overlap*

To get a better understanding of how individual cells contribute to the entirety of the behavioral task, each cell was plotted across time for all of the relevant behavioral factors. Any time a cell reached significance for any of the relevant regression factors within a 200ms time window, a different color was plotted based on the factor of significance. This way, each cell could be followed throughout the task, and the length of encoding as well as the actual factor that was encoded could be seen. Finally, a direct comparison of cells in different behavioral events, reaching significance for different regression factors was compared using a Venn diagram, to better understand how individual cells overlapped in time and encoding. The first Venn Diagram represents Action-Goal-Outcome encoding during the events Nose Side In (action) and Reward Port In (Goal and Outcome). The second Venn Diagram represents Goal-Outcome-Integration encoding at 3 different time points around Reward Port In (see methods). Each region is plotted separately, with the 3-factor regression shown for that region alone. A separate analysis was done for Action-Goal-Cue (Figures 15-19), and Previous Outcome-Goal-Cue (Figures 20-24).

Difference in encoding schemes across events and between cells demonstrates other regional differences. When comparing Action cells to cells that encode Goal or Outcome, both the DMS and Shell show the greatest separation of cells, having very few cells overlap their encoding during different behavioral events (Figure 16C and Figure 18C). On the other hand, the DLS and Core show more overlap within the same cell across behavioral events and different encoding factors (Figure 15C and Figure 17C). In

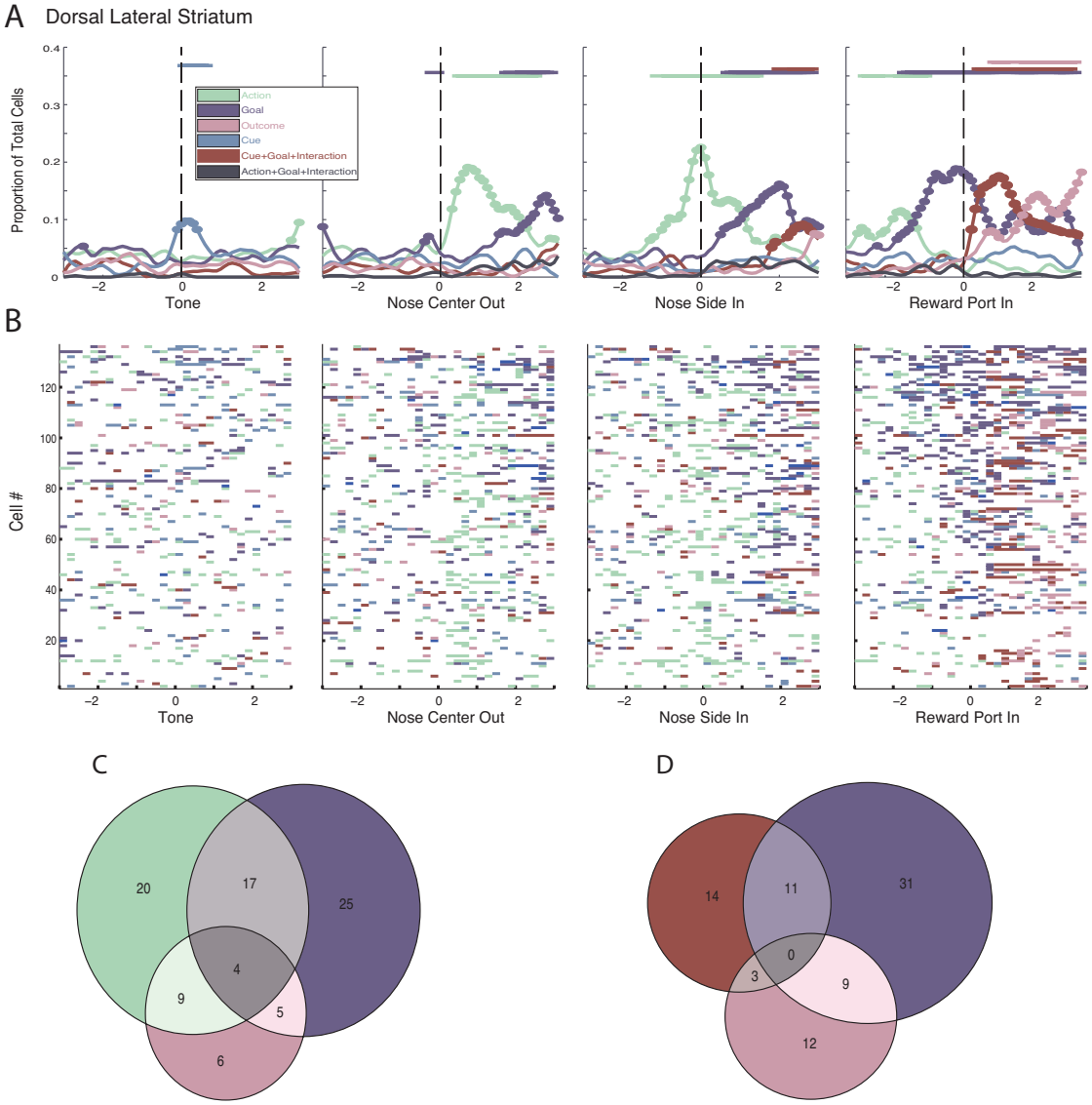
fact, only 6 of the 32 cells in the Core only encode the factor Action, while 14 cells overlap between Action and Goal, and 5 overlap with Action, Goal and Outcome (Figure 17C and 17D). The OFC has the highest amount of integration of all 5 brain regions, showing that the OFC encodes a little bit of everything across all behavioral events, without having one distinct role (Figure 19C and 19D).

When comparing how cells encode information at the reward port, the DLS again stands out as the region that has the highest number of cells that only integrate information about the Cue, the Goal and the Outcome (Figure 15D), while still a large proportion of cells that contribute to this integration also showed significance for the Goal. The DMS has the fewest number of cells that overlap between factors. The OFC and Shell show almost no overlap between cells that encode 'Outcome' and those that are significant for Cue, Goal and Outcome (Figure 16D). This could mean that the cells that integrate all 3 signals are doing so for the motor movement, while a separate signal encodes whether or not a reward was received.

The DMS shows relatively even encoding across all factors and all events. The DLS shows a brief increase in firing for the Cue, for the Action, and then has a much longer responses to the reward related factors, with a nearly even distribution between the Goal, the outcome, and their integration. The Core shows very brief responses for the Action, very brief responses for the Goal, and longer responses in cells that fire for Outcome. The Shell has cells that fire for a longer period of time for the Cue, with short periods of encoding for the Action. At Reward port in, responses to the outcome are the longest. The general firing rate patterns of the DMS and Shell are similar, while the firing rate patterns of the DLS and Core are similar. Although it is not clear from this study how

these regions are acting differently for model-free vs. model-based encoding, it does appear that a the schematic in Figure 2B, at least as far as encoding patterns go, gives support to a DMS-Shell relationship, and a DLS-Core relationship.

Figure 15: Dorsal Lateral Striatum



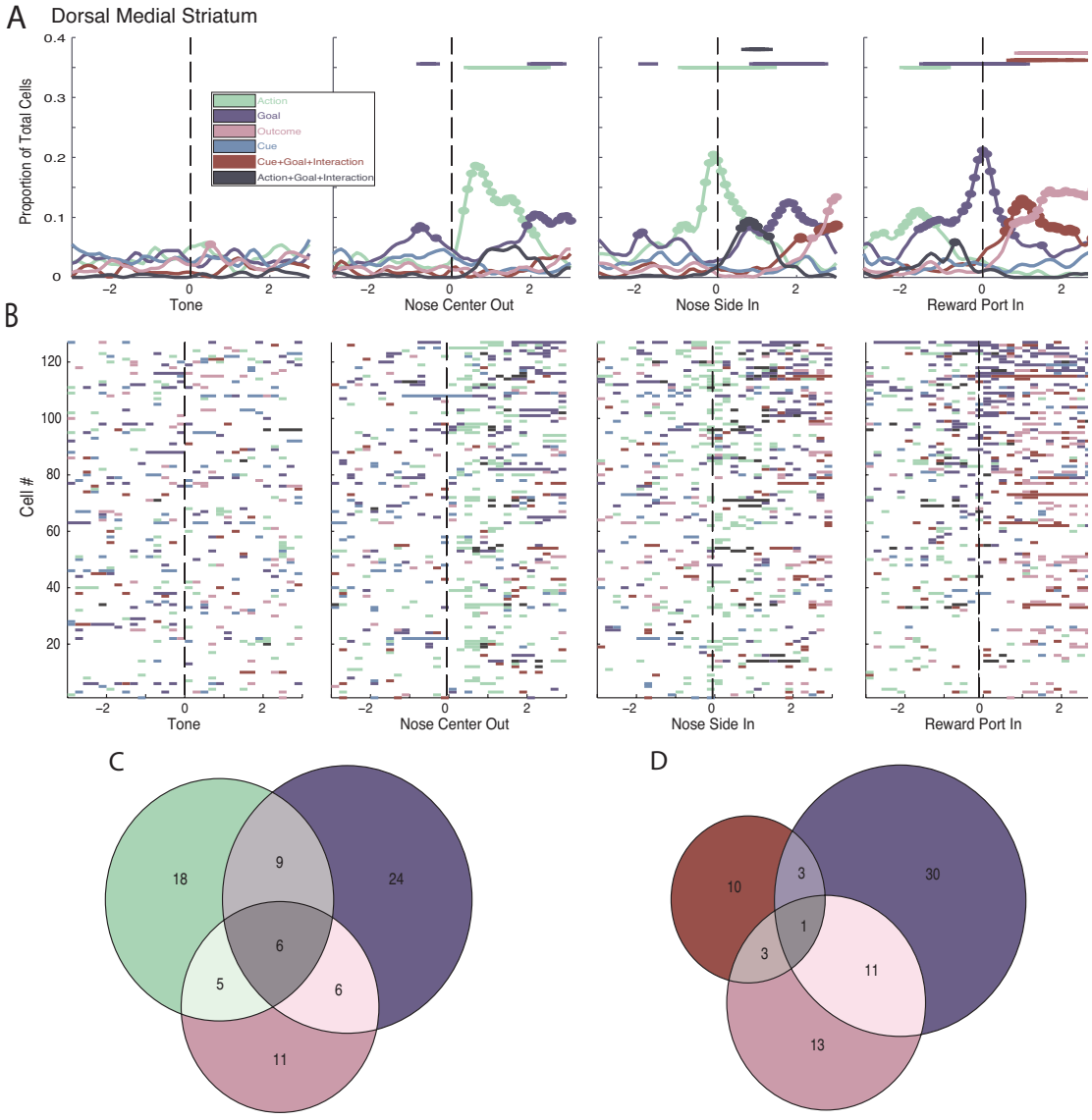
A: Encoding by Region. The 3-Factor regression Action-Goal-Cue is plotted for all of the relevant factors for the cells in just the DLS, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the DLS is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Goal' encoding at reward port in, and this sorting order remained consistent across all behavioral events.

C: Event Overlap. The Venn Diagram represents Action-Goal-Outcome encoding during the events Nose Side In (-0.5 to 0.5 seconds for action), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

D: Event Overlap. The Venn Diagram represents Goal-Outcome-Integration encoding at 3 different time points around Reward Port In (Goal: -0.5 to 0.5 seconds; Outcome: 0.5 to 1.5 seconds; Integration: 1.5 to 2.5 seconds).

Figure 16: Dorsal Medial Striatum



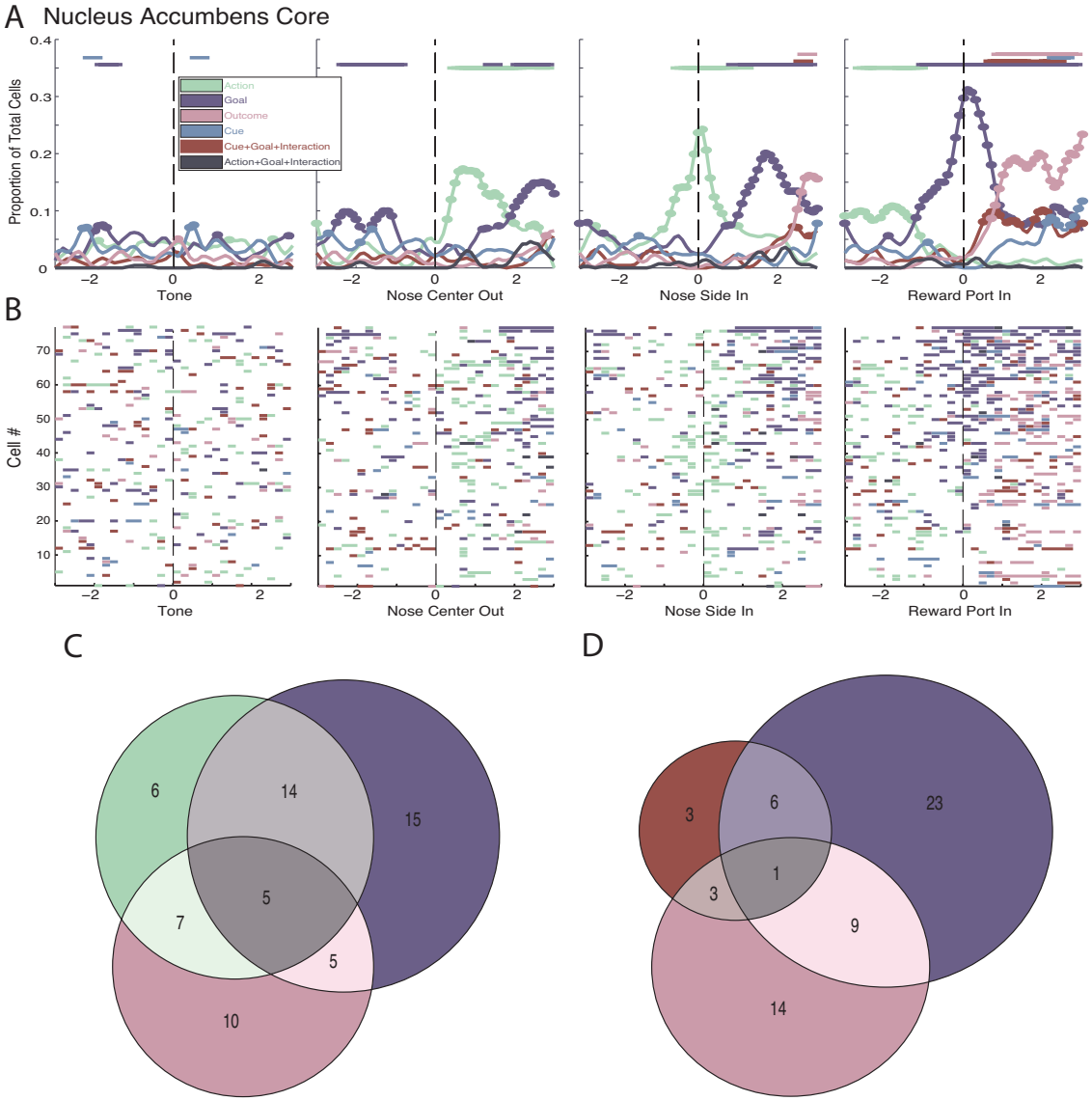
A: Encoding by Region. The 3-Factor regression Action-Goal-Cue is plotted for all of the relevant factors for the cells in just the DMS, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the DMS is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Goal' encoding at reward port in, and this sorting order remained consistent across all behavioral events.

C: Event Overlap. The Venn Diagram represents Action-Goal-Outcome encoding during the events Nose Side In (-0.5 to 0.5 seconds for action), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

D: Event Overlap. The Venn Diagram represents Goal-Outcome-Integration encoding at 3 different time points around Reward Port In (Goal: -0.5 to 0.5 seconds; Outcome: 0.5 to 1.5 seconds; Integration: 1.5 to 2.5 seconds).

Figure 17: Nucleus Accumbens Core



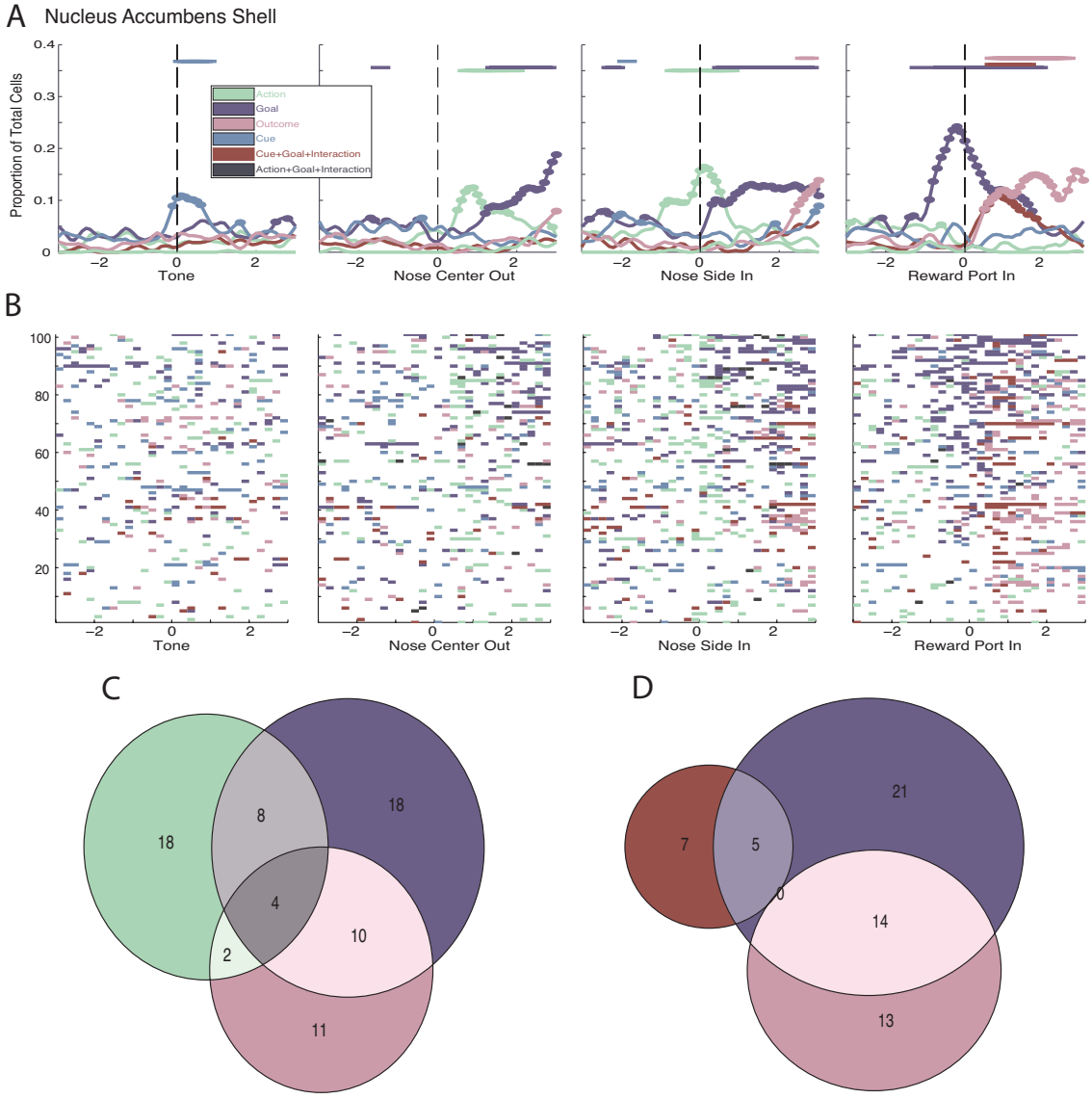
A: Encoding by Region. The 3-Factor regression Action-Goal-Cue is plotted for all of the relevant factors for the cells in just the Core, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the Core is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Goal' encoding at reward port in, and this sorting order remained consistent across all behavioral events.

C: Event Overlap. The Venn Diagram represents Action-Goal-Outcome encoding during the events Nose Side In (-0.5 to 0.5 seconds for action), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

D: Event Overlap. The Venn Diagram represents Goal-Outcome-Integration encoding at 3 different time points around Reward Port In (Goal: -0.5 to 0.5 seconds; Outcome: 0.5 to 1.5 seconds; Integration: 1.5 to 2.5 seconds).

Figure 18: Nucleus Accumbens Shell



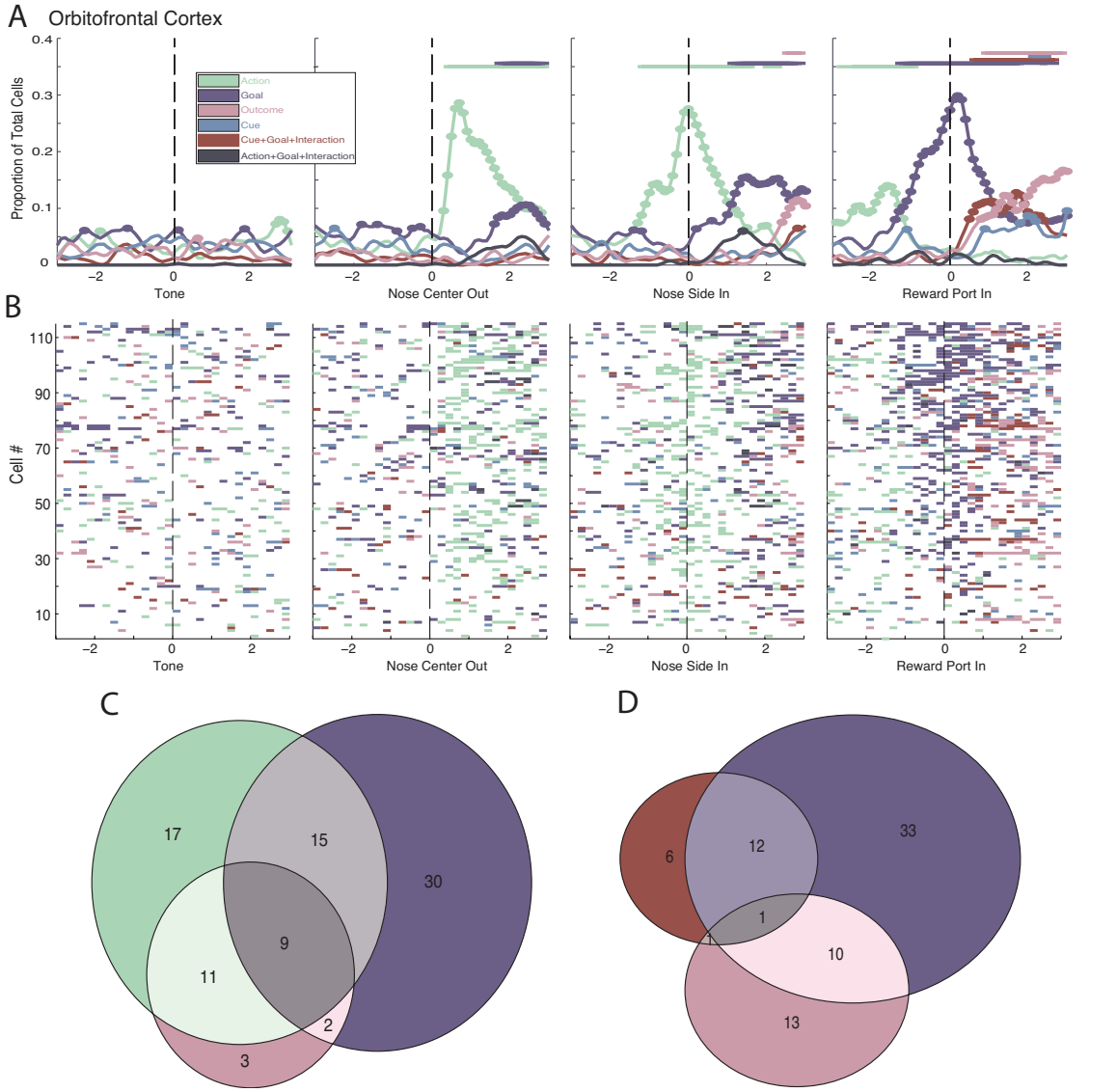
A: Encoding by Region. The 3-Factor regression Action-Goal-Cue is plotted for all of the relevant factors for the cells in just the Shell, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the Shell is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Goal' encoding at reward port in, and this sorting order remained consistent across all behavioral events.

C: Event Overlap. The Venn Diagram represents Action-Goal-Outcome encoding during the events Nose Side In (-0.5 to 0.5 seconds for action), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

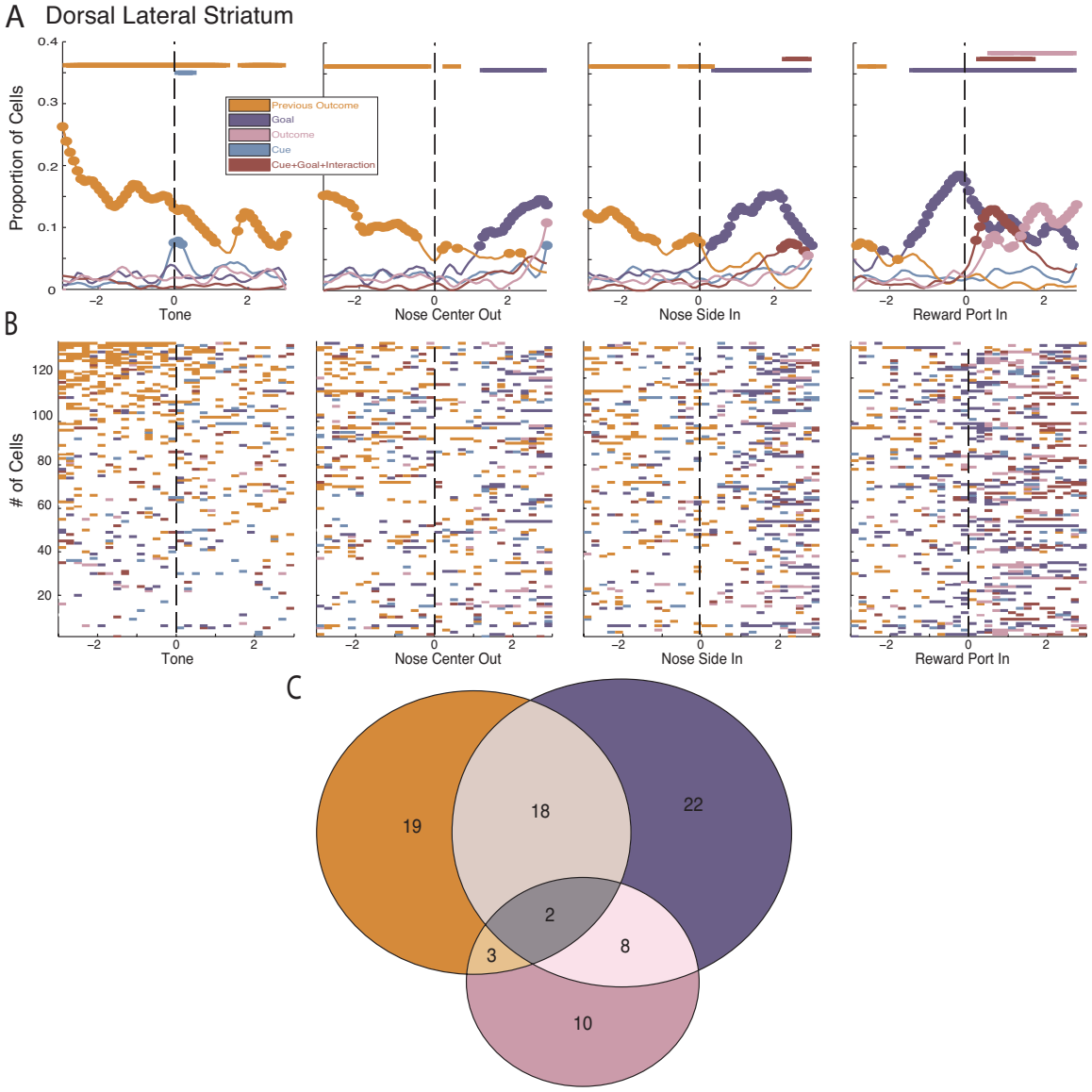
D: Event Overlap. The Venn Diagram represents Goal-Outcome-Integration encoding at 3 different time points around Reward Port In (Goal: -0.5 to 0.5 seconds; Outcome: 0.5 to 1.5 seconds; Integration: 1.5 to 2.5 seconds).

Figure 19: Orbitofrontal Cortex



A: Encoding by Region. The 3-Factor regression Action-Goal-Cue is plotted for all of the relevant factors for the cells in just the OFC, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
B: Individual Cell Plots. Each cell that was recorded from in the OFC is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Goal' encoding at reward port in, and this sorting order remained consistent across all behavioral events.
C: Event Overlap. The Venn Diagram represents Action-Goal-Outcome encoding during the events Nose Side In (-0.5 to 0.5 seconds for action), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).
D: Event Overlap. The Venn Diagram represents Goal-Outcome-Integration encoding at 3 different time points around Reward Port In (Goal: -0.5 to 0.5 seconds; Outcome: 0.5 to 1.5 seconds; Integration: 1.5 to 2.5 seconds).

Figure 20: Dorsal Lateral Striatum

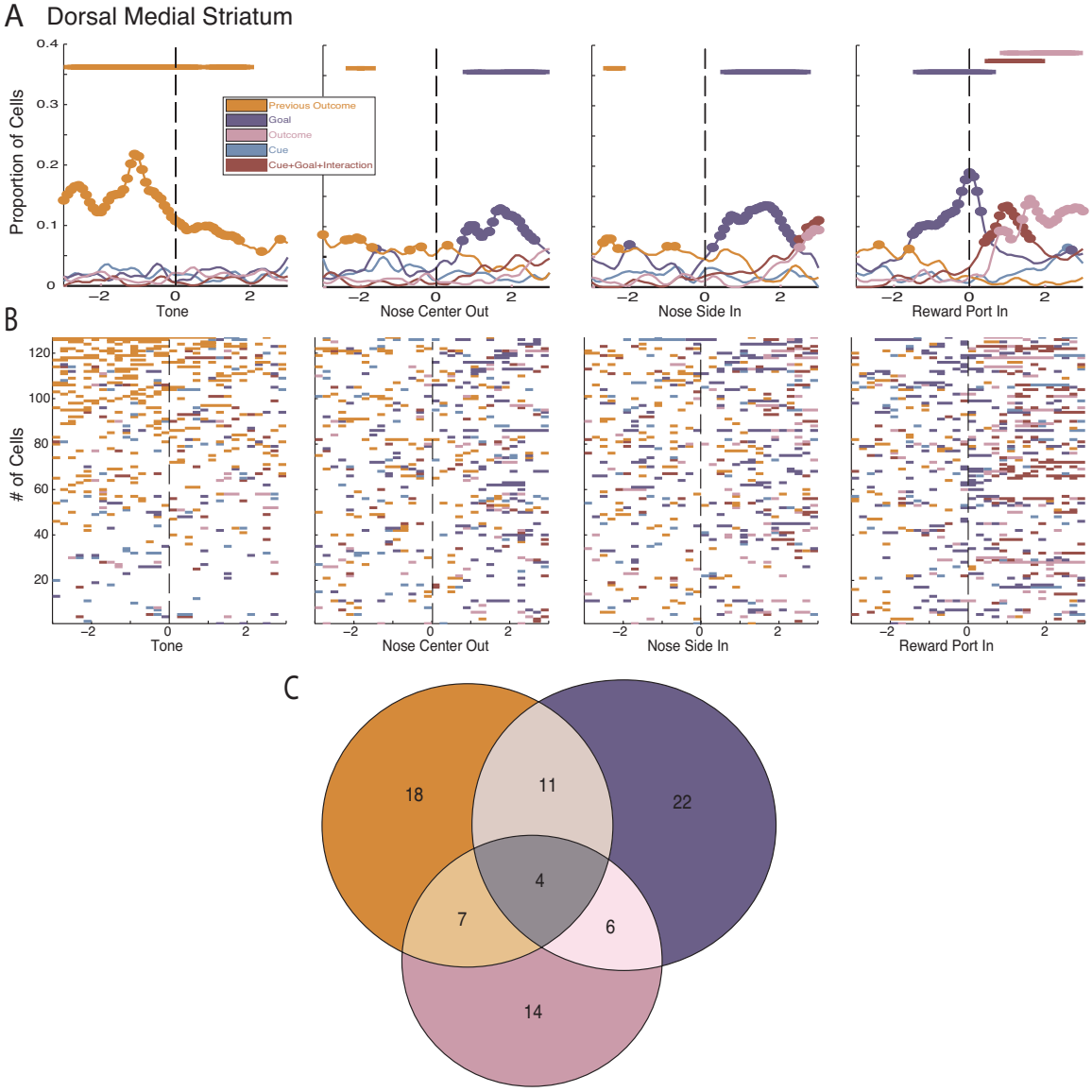


A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the DLS, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the DLS is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.

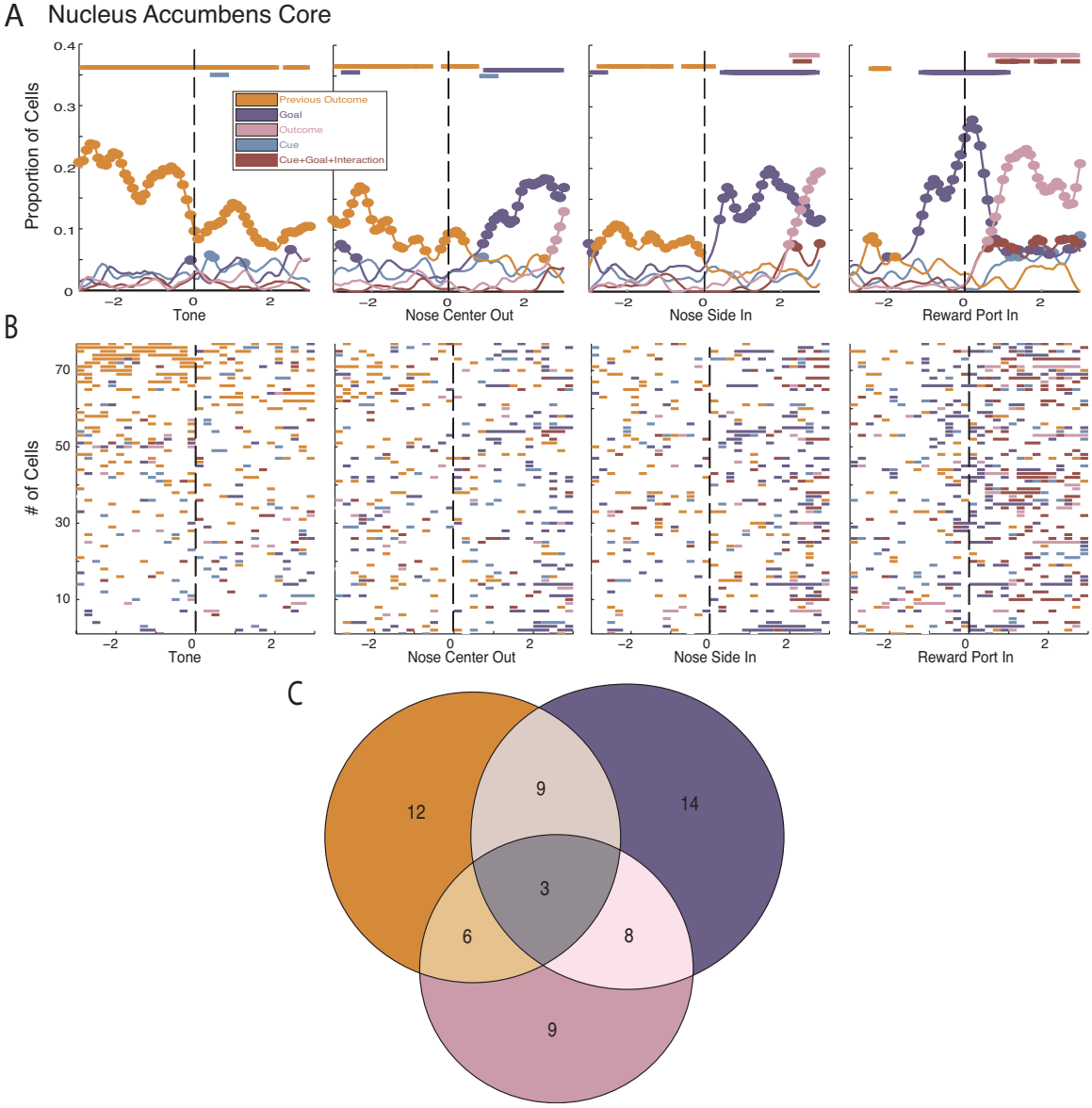
C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Figure 21: Dorsal Medial Striatum



A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the DMS, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.
B: Individual Cell Plots. Each cell that was recorded from in the DMS is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.
C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Figure 22: Nucleus Accumbens Core

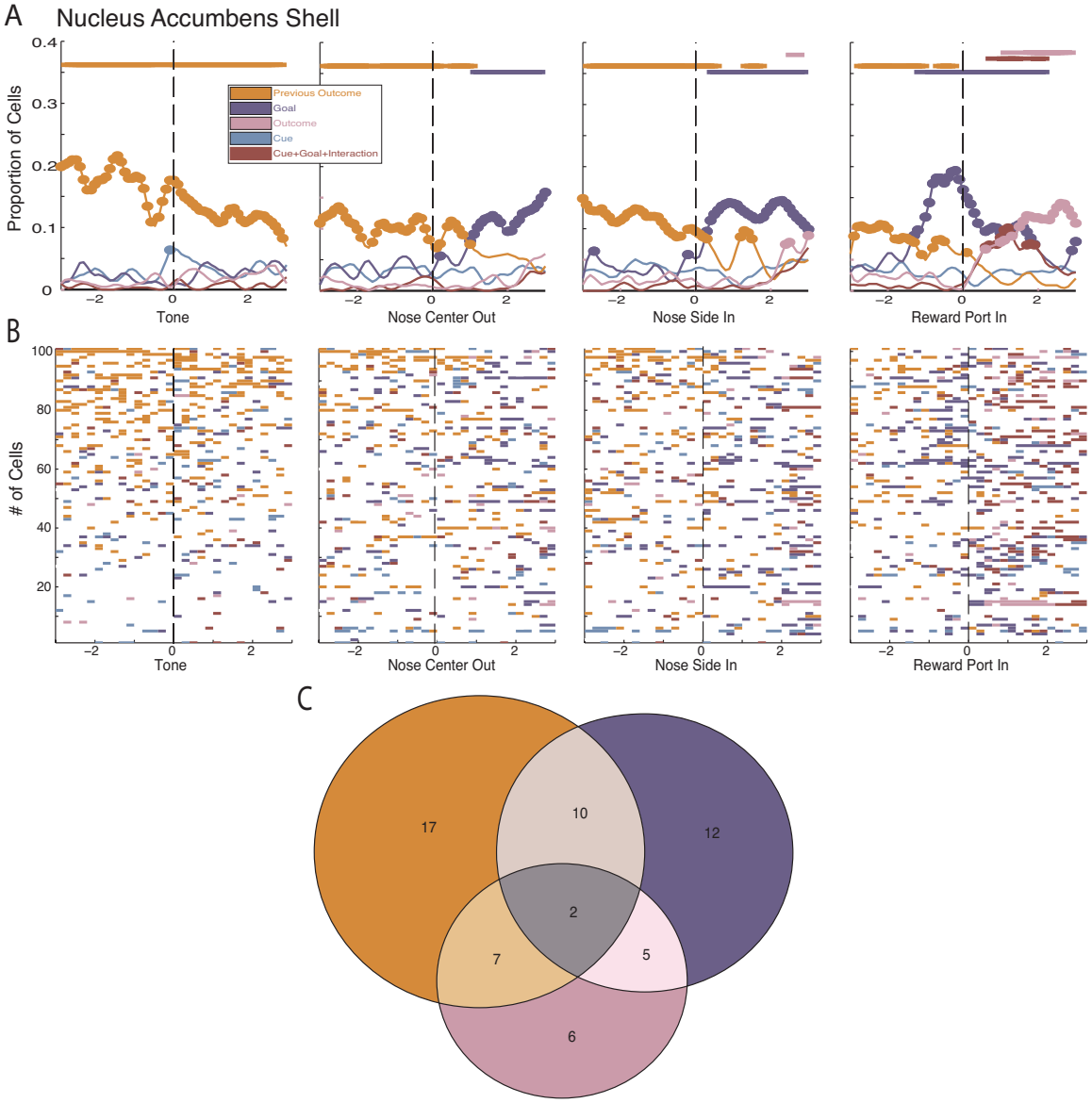


A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the Core, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the Core is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.

C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Figure 23: Nucleus Accumbens Shell

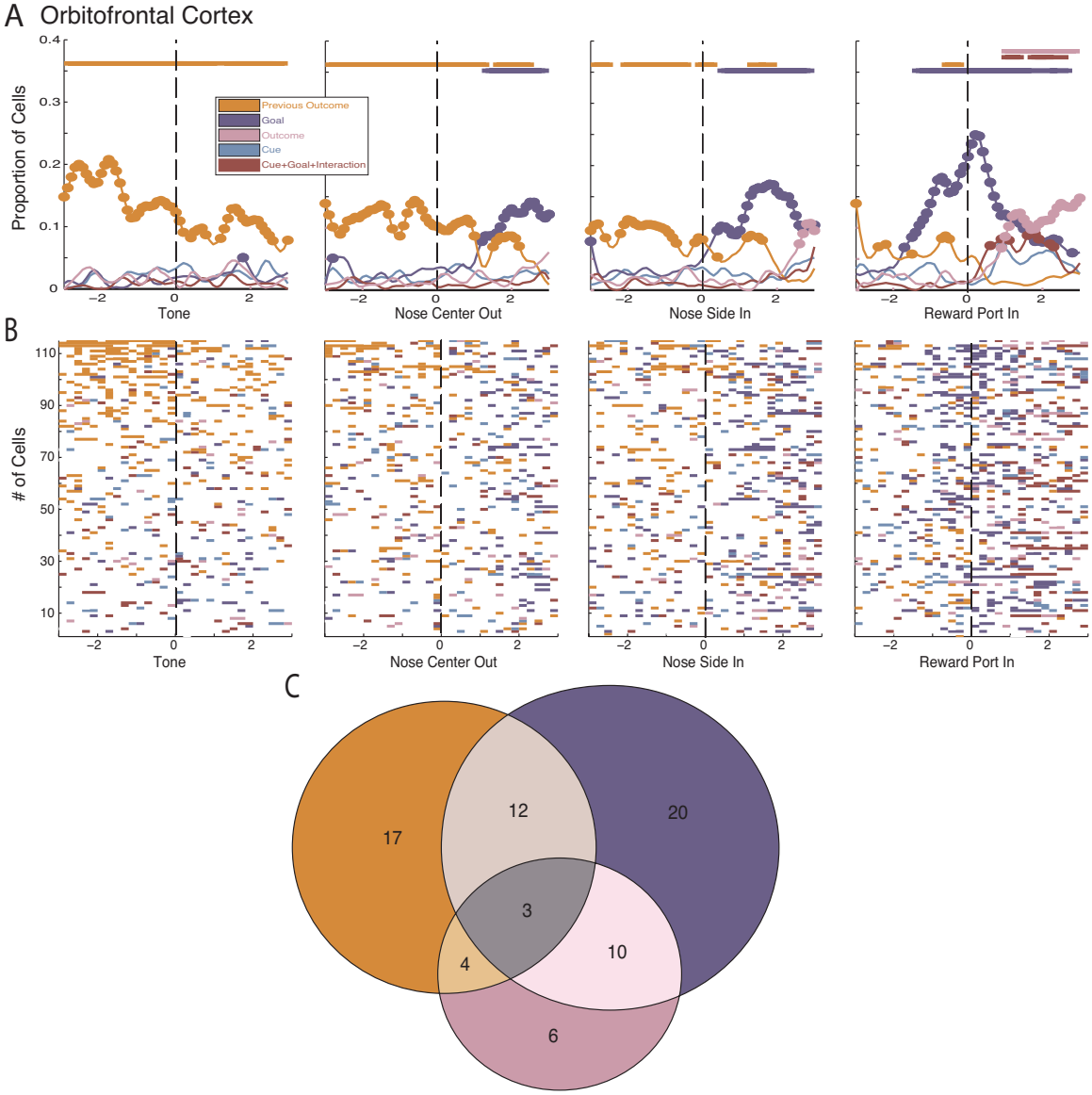


A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the Shell, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the Shell is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.

C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Figure 24: Orbitofrontal Cortex



A: Encoding by Region. The 3-Factor regression Previous Outcome-Goal-Cue is plotted for all of the relevant factors for the cells in just the OFC, in 6 second time windows surrounding the relevant behavioral events. The total proportion of cells that reach significance are shown, and the filled circles represent when a region reached significance (binomial test, $p < 0.05$), while the corresponding bar across the top indicates when 2 adjacent bins reach significance. Bin width of 500ms, sliding in 100ms steps.

B: Individual Cell Plots. Each cell that was recorded from in the OFC is plotted individually. When a cell reached significance for any of the relevant multiple regression factors, a colored line is used to indicate the time window (200ms, non-overlapping) of significance for that factor. In this way, all of the different regression factors are visualized for each cell, across all events. The individual cells were sorted for the amount of 'Previous Outcome' encoding at the Tone, and this sorting order remained consistent across all behavioral events.

C: Event Overlap. The Venn Diagram represents Previous Outcome-Goal-Outcome encoding during the events Tone (-2 to -1 seconds for Previous Outcome), Reward Port In (-0.5 to 0.5 seconds for Goal and 0.5 to 1.5 seconds for Outcome).

Chapter 4: Final Conclusions

In task after task, manipulation after manipulation, specific regions are thought to behave in a specialized manner, which is clearly delineated along anatomical borders. With that in mind, the expectations going into this experiment were to find very clear patterns of activity that would strictly segregate according to region. The fact that this behavioral task so strongly separated all aspects of a reward based decision-making task contributed to this hypothesis. If Cue can be so distinctly separated from Action- in both time and meaning- which is also separated from the Goal, then these aspects of decision-making that are so clearly seen in other lesion and electrophysiology studies should be even more obvious in the current study.

Universal Action Encoding

The strongest and most overwhelming result of this study is that the individual brain regions are highly redundant. Surprisingly, the most universal and consistent coding seen in all 5 brain regions was for 'Action'- a differentiation between contralateral and ipsilateral movements. Importantly, this study specifically disentangled actions from the upcoming reward, by having the cue that signaled direction separate and unattached to

the cue that signaled the reward. The rat used a visually guided cue to determine movement direction, and then proceeded to the reward port after successfully completing the nose poke action. Contrary to the hypothesis, all 5 brain regions encoded action, independently of reward in a large fraction of cells (ranging from 17-33% of task-related cells). Even more surprisingly, the OFC had the highest proportion of cells encoding the action variable, at 34%. The Shell had the lowest proportion of cells encoding action, and also did so for the shortest amount of time.

Importantly, this behavioral task as well as the analysis, was very restrictive of how to encode the action variable. The Cue that began each trial (a tone) only signaled the upcoming reward (food vs. water). There was no pre-motor signal that gave the rat any information about what movement he would make. Therefore, as expected, there was no pre-motor movement related activity. This is different than many other tasks that often impose a delay between a reward/movement instructive cue, and a trigger stimulus that then tells the animal to go (Gage et al., 2010; Apicella et al., 1992; Hollerman et al., 1998; Pasupathy and Miller, 2005; Samejima et al., 2005; Hikosaka and Watanabe, 2000).

Additionally, when examining specific motor activity, the analysis was restricted to cells that only showed specific directional selectivity- so more firing on a contralateral trial vs. an ipsilateral trial. This analysis excluded cells that had a change in firing rate over baseline, but did not encode a specific direction. To further fine-tune the analysis, the multiple regression was restricted to cells that showed ONLY a main effect of 'Action'. If a cell encoded another aspect of the task at the same time as it became significant for Action, it was not included in these Action-Only results. This is unlike

many other studies that compare activity within a specific time window to baseline, without necessarily differentiating different movements. It also excludes cells from this particular analysis that encode something other than a purely movement related response.

Because of this highly restrictive analysis, it was even more surprising to find such abundant encoding of contralateral vs. ipsilateral movements, with nearly the same temporal profile, in all 5 brain regions. Studies that have examined strict motor firing in the basal ganglia have often found it to be very context dependent (Hikosaka et al., 1989b; Rolls et al., 1983; Tremblay et al., 1998), so the fact that we still found such abundant motor movement, with differential starting points and distinctly separate reward-value predictions, is a significant contribution to understanding the basal ganglia literature. Many decision-making tasks overlay the combination of action + value to differentiate neural firing rate patterns, and cannot speak to the separate contributions within each subregion. This task can.

Dorsal Striatum: Directional coding in both the medial and lateral dorsal striatum has been seen in a variety of tasks, throughout the reward and learning literature (Wiener, 1993); (Yeshenko et al., 2004; Stalnaker et al., 2010; Van Der Meer et al., 2010; Thorn et al., 2010; Barnes et al., 2011; Schmitzer-Torbert and Redish, 2008; Schmitzer-Torbert and Redish, 2004; Tang et al., 2007; Chang et al., 2002; Gage et al., 2010; Bryden et al., 2012). The basal ganglia has long been thought to be an area important to the selection of actions (Mink, 1996; Redgrave et al., 1999; Samejima and Doya, 2007). Lesion studies have also demonstrated the role of the dorsal striatum in executing specific directional

responses (Brasted et al., 1997; Carli et al., 1989; Cook and Kesner, 1988; Brown and Robbins, 1989; Döbrössy and Dunnett, 1997).

Based on recent lesion literature, there is a strong consensus on the medial-lateral divide in the dorsal striatum, and the subsequent role each has in reward related behavior and learning (Yin and Knowlton, 2006; Balleine et al., 2007; Balleine and O'Doherty, 2010; Yin et al., 2008, White, 2009; Redgrave et al., 2010; Devan et al., 2011). It is therefore somewhat surprising to find a roughly equal contribution of neurons from DMS and DLS to movement-specific aspects of my task. In contrast to the results seen in most lesion studies, others who used single unit recordings to directly compare DMS and DLS found similar results to this study. In tasks directly testing the contributions of DMS and DLS to habit vs. goal-directed behaviors, recordings done in the rodent striatum have all found surprising overlap between the two regions in the amount of direction encoding (Kim et al., 2013; Kimchi et al., 2009; Stalnaker et al., 2010; Thorn et al., 2010).

Despite similar proportions of neurons encoding Action between regions, there were differences between the DMS and DLS within this group of neurons. To further examine how the individual units contribute to the regional encoding of Action, we examined the individual firing rate patterns to create an index that tracks the trial type with the highest firing rate. That is, if a cell differentiates between contra vs. ipsi movements, which of those two trials had the higher firing rate. From this index, the proportion of cells of each trial type was plotted over time, and each bin was tested for significance ($p < 0.05$, binomial test). At the time of peak Action encoding, which happens between 400ms to 800ms after nose center out, the DMS shows a significant regional bias for the contralateral movement direction. In contrast, the DLS shows no

regional bias. This means that of the cells that show a significant contra/ipsi differentiation, those in the DMS are firing more for the contralateral action, while those in the DLS are firing equally for contra and ipsi movements (Figure 7C).

This difference between DMS and DLS contribution to Action encoding is somewhat surprising. Lesion studies in the rodent have often found varying forms of a contralateral deficit in both areas of the striatum (Döbrössy and Dunnett, 1997; Carli et al., 1989; Brasted et al., 1997; Brown and Robbins, 1989). However, electrophysiology studies examining either or both areas have led to conflicting results. Some studies show no bias in either region, including a study from the Berke lab (Gage et al., 2010; Stalnaker et al., 2010). In maze tasks, although movement/turn related activity was encoded in areas of the striatum, direction specific bias was not reported in studies from the other labs as well (Berke and Eichenbaum, 2009; Schmitzer-Torbert and Redish, 2004, 2008; van der Meer et al., 2010). In contrast to these results, the Graybiel lab (Thorn et al., 2010), did find directional bias in both DMS and DLS, with a stronger contralateral bias in the DLS. However, a nose poke task, more similar to the one employed in this study, that was only examining the DMS found an increase in firing rate in the DMS for contralateral turns (Bryden et al., 2012).

A possible explanation for there were different regional biases between the DMS compared to the DLS could be due to the number of FSIs. In the DLS, 27% of the task related responses came from FSIs. In the DMS, only 10% of the task related responses during the action event were from FSIs. It could have been that since there are relatively more FSIs in the DLS, then the stronger contribution of FSIs to the recording database served to balance out the distribution of bias between contra and ipsi movements, as was

seen in Gage (Gage et al., 2010), MSN-FSI pairs always showed an opposing bias. However, when examining either just MSNs or just FSIs, the same patterns are found- that the DLS has no specific directional bias while the DMS is biased in the contralateral direction.

Another possible explanation for the difference in regional bias between the DMS and DLS could be due to the actual mode of response. Even though the DLS is thought of as the sensorimotor striatum, due to its anatomical connections and task-related responses, it is possible that single neuronal contributions to movement are in a more big picture, broad sweeping variety- so that individual contralateral and ipsilateral directionality in a small nose poke task may not be enough to engage specific biases to either direction. Carelli (Carelli and West, 1991) demonstrated that 38% of DLS neurons responded to whole body movements, and that there were some neurons still tuned to specific directions. In contrast, the DMS may be more concerned with specific response-related movements that cause a greater increase in the contralateral movement direction due to downstream connections to the substantia nigra.

The sensorimotor striatum receives convergent projections from the somatosensory cortex, primary motor cortex and premotor cortex (Haber et al 1994 (Haber, 2003; Parent and Hazrati, 1995). In recording studies, specific MSNs have been found that fire in relation to sensorimotor activity of single body parts (Alexander and DeLong, 1985; Carelli and West, 1991), and from the connections of the DLS to the GPi/SNr to the thalamus, they then project back to the motor cortical area (Parent and Hazrati, 1995; Alexander et al., 1986; Haber et al., 2000). Due to these connections to a more broad range of motor cortex, it is likely that the DLS is only concerned with the

broad movements of contra vs. ipsilateral directions, activating similar numbers of direct and indirect striatal pathways, while the DMS, which may be more intricately involved in specific movements, fires more for the contralateral space due to the requirement of activating more direct neurons.

Nucleus Accumbens: The original hypothesis did not anticipate a large amount of action-specific encoding in the Nucleus Accumbens. Since the NAC is usually thought of the limbic-motor interface (Mogenson et al., 1980), due to its anatomical connections (Groenewegen and Russchen, 1984; Heimer et al., 1991; Wright and Groenewegen, 1995); Voorn et al., 2004), the hypothesis was that the NAc would integrate Action + Value, or Action + Cue, rather than simply fire for Action alone. In fact, previous lesion work in the NAc found no deficit in a spatial task (Kelley et al., 2005; Kelley, 1999).

Similar to the different results seen when contrasting single unit studies vs. lesion studies in the Dorsal Striatum, studies that examine single units within the NAc have found action-related encoding in rodents (Taha et al., 2007; Ito and Doya, 2009; Chang et al., 2002; Kim et al., 2009; Roesch et al., 2009). Most of these rodent studies examining neural activity have focused their analysis on how the movement related activity signals upcoming reward choices, rather than specifically examining the movement themselves. Additionally, when examining movements, prior studies have just examined the general movement-related time period compared to baseline, rather than having a task designed specifically to examine contralateral vs. ipsilateral movements.

Contrary to the original hypothesis, we found a significant amount of encoding for specific contralateral vs. ipsilateral movements in both regions of the Nucleus

Accumbens. The Shell had the lowest proportion of cells encoding Action, reaching a peak of 18%, while the Core had a proportion of 25%. Both regions began encoding movement as the rat began moving, although the Shell reached significance at a later time point than the other 4 regions. These results were for cells that specifically differentiated contralateral and ipsilateral movements- this was not a general movement-related activity pattern, unlike other studies (Figure 7A).

An additional result that is still somewhat puzzling is the fact that within the cells that encoded action, there were significant proportions of cells with contralateral directional bias in both the Core and Shell (Figure 7C). The Shell only briefly (but significantly) responded with a contralateral bias for 600ms, right at peak encoding. The Core bias began towards the end of peak encoding, and lasted as long as the Shell. In previous single unit recording studies that examined directional encoding in the NAc, no other study has reported a significant bias, typically finding similar proportions of contra and ipsi favoring neurons (Taha et al., 2007; Ito and Doya, 2009; Kim et al., 2009; Roesch et al., 2009).

It is possible that there is a motor component involved in the NAc firing of a subset of cells that explains why there is a brief contralateral bias in the Core and Shell. McGinty (McGinty et al., 2013) examined the firing rates of NAC neurons in a task that allowed the animals to flexibly approach a nose poke, and found that cue-evoked excitation predicted the movement initiation latency and speed of the flexible approach response, as well as proximity to the target. The task event that we analyzed here wasn't under the same flexible approach paradigm that was used in the McGinty task- the movements still had a specific beginning and end location- but the speed and vigor with

which the rat made his movements, and the flexibility of left/right not being signaled ahead of time, could have been encoded in these neurons, and reflected in the subsequent firing rate bias. However, a specific analysis examining speed and vigor is beyond the scope of the current thesis project.

Their results, as well as the results in the current study, can possibly be explained by the fact that different groups of neurons encode specific actions, while at another time in the task, other groups of neurons in the NAC encode the upcoming goal choice. Most previous research, even those involving effort or long delays, still do not have the drawn out behavioral events seen in the current task. Additionally, since Cue, Action and Goal were separated in both meaning and time, it gave ample opportunity to dissect what each brain region was doing. From these results, it is apparent that in both Core and Shell, there are subgroups of neurons that specifically encode directional movements in the pursuit of reward, without having those movements be directly tied to reward.

Orbitofrontal Cortex: Even more surprising than the significant amount of Action encoding seen in the NAC, as well as the strong contralateral bias in that region, was the fact that the region with the highest proportion of Action encoding, in the absence of reward value, was the OFC. Primate research in the OFC has long demonstrated no directional bias in a wide range of single unit studies (Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006; Kennerley and Wallis, 2009a; Abe and Lee, 2011). Contrary to findings in non-human primates, most single unit studies in rodents have shown directional selectivity in the OFC (Feierstein et al., 2006; Furuyashiki et al., 2008; Young and Shapiro, 2011; Roesch et al., 2006). Only recently has there been an

exception in primate literature, in a study by Luk and Wallis (Luk and Wallis, 2013), as well as one other primate study by Tsujimoto (Tsujimoto et al., 2009).

In trying to differentiate the strong difference in findings of movement encoding between the primate and rodent literature, the answer may be in the specific motor component of the behavioral task. A majority of primate studies that have not found directional selectivity in the OFC used tasks that involved eye movements (saccades) (Padoa-Schioppa and Assad, 2006; Wallis and Miller, 2003), or lever reaching (Tremblay and Schultz, 2000), and only after performing those movements did the rat than indirectly receive a reward- he did not actually execute a motor pattern to get the reward himself. It is possible that the rat studies, which involved the rats physically moving through space to specific, spatially distinct goals (reward locations), caused the difference in firing rate for movement, seen in the rodent OFC. In contrast to the primate OFC, studies examining primate *striatum* often show directional selectivity, but this may be due to the specific anatomical connections in the primate caudate to the visual cortex, that is specific to the saccade movement.

Another possible explanation, as brought up in Luk and Wallis (Luk and Wallis, 2013), is that previous primate OFC studies, despite utilizing an instrumental response, may have actually been solved using a Pavlovian mechanism- that is, the pattern of stimulus and outcome was so well learned that it no longer relied on planning the actual motor movement. Like the Luk and Wallis task in primates, the current task also did not require the rats to make the exact same movement to get a reward, preventing a long term mapping of precise movements with reward. This idea can also be thought of as disentangling actions and rewards.

Even though the findings in this study about Action selectivity in the OFC correspond to other rodent electrophysiology studies, the major difference is that there is also a significant directional bias in the OFC. Similar to the NAc and DMS, the OFC initially showed a strong contralateral bias as the rat moved from the center nose port. In contrast to the other regions, the OFC then switched to an ipsilateral bias at 2.2 seconds after nose center out. This was as encoding of action began tapering off, so it was with a smaller group of cells.

One of the main studies to find directional selectivity in the rodent OFC was from the Mainen group (Feierstein et al., 2006). They recorded from a similar area in the rodent OFC and had a similar task, where direction was not signaled prior to movement. They found that during the response period, 41% of the cells were directionally selective, a very similar proportion to what was found in the current results. Even during the entry into goal port, 35% of cells were still selective for the left/right location of the ports.

Research in the OFC that has found directional selectivity has not found a bias for either ipsilateral or contralateral movements (Feierstein et al., 2006; Furuyashiki et al., 2008; Young and Shapiro, 2011; Roesch et al., 2006). In the OFC, Feierstein did find that speed of movement modulated the activity of some direction-selective neurons, which could explain the time course of selectivity in the current data- that rats were faster for contralateral movements, causing a bias early in the event, then the slower, ipsilateral cells demonstrated their bias later. Additionally (Furuyashiki et al., 2008) did not find a specific bias in contra vs. ipsi movements, but did find that the rats moved faster for right trials than left trials. They found that action-related neuronal firing rates were correlated with speed of movement. A more in depth analysis into specific timing-related firing rate

differences is beyond the scope and objective of this thesis, however it would be an interesting avenue for further research.

Interestingly, when looking at directional bias on the days when the rat was only sated or restricted, the switch in contra/ipsi bias no longer is apparent. Directional selectivity is still highly significant in all 5 regions, but there is also only a small significance in the OFC bias to the contralateral side, and it happens later than on the food and water restricted days (Figure 8A, inset). This could mean that there is an invigoration component to the firing rate differences seen in the direction, that creates a bias when the rats are in a more motivated sate.

Allocentric vs. Egocentric encoding

Previous work examining the role of the striatum in reward related tasks has often questioned the idea that cells in the DMS and DLS may be encoding space, rather than movement (Mizumori et al., 2009). To account for this potential explanation, we used a modified 5-poke chamber, and utilized the center 3 nose ports as starting points, randomized across trials. The design of the nose poke parameters ensured that the starting central nose port changed randomly, so that each of 3 different central nose pokes was utilized throughout each recording session. This way, each movement was made to a different allocentric location- the specific movements were in relationship to the rat's egocentric space, rather than specific port entries. In this way, encoding of spatial vs. allocentric movements were controlled for. This also makes this task unique compared to most other rodent reward paradigms, who used the same central nose pokes, and left vs. right fluid wells to deliver reward (Roesch et al., 2009).

The contralateral and ipsilateral movements made in this task were to egocentric space. The analysis of direction performed in the multiple regression was used based on the head movement made, not the starting port. If starting port were more highly significant than the movement, than the Action-Only factor would fail to reach significance.

Despite being able to determine if these firing rate changes were allocentric vs. egocentric, it was not possible to determine if they were turn-specific, or head direction firing rate changes, since the movements made by the rats involved moving the head in separate directions, but each movement was relatively stereotyped. There was no comparison for movements made away from the nose poke vs. those done at the nose pokes because there was no way to precisely measure turning behavior away from the laser-guided nose poke holes. Additionally, there were no EMG recordings taken to determine if the neck related movements completely corresponded to the neural firing rates seen during action.

A more in depth video analysis with corresponding EMG recordings would probably help explain the differences seen in encoding in the DMS vs. DLS, as a difference in general body orientation vs. specific head movements. If the DMS is more concerned with fine tune movements, than the subtle contralateral and ipsilateral movements may engage more of the direct pathway within the basal ganglia, and a subsequent increase in firing rate for contralateral movements. If the DLS is more whole-body oriented, than the entirety of the direct and indirect pathways may be involved in orienting towards the adjacent nose pokes, and cause an overall split between firing more for contra vs. ipsi turns.

Timing of Action

Other results from the striatum have found post-movement encoding, or encoding only after response initiation (Lau and Glimcher, 2007; Kim et al., 2009; Kimchi and Laubach, 2009a) but the results from this study show involvement of all 5 brain regions as the rat begins movement into the adjacent, lit nose port (Figure 7A). There is some post-movement activity that continues in all 5 brain regions up to 1.5seconds after completing the nose poke movements.

In primates and some rodent studies, action coding in striatal regions is often found to encode reward expectations in the upcoming motor movement (Hikosaka et al., 1989a; Apicella et al., 1992; Hollerman et al., 1998). However, this particular behavioral task did not give any pre-instruction motor related information, and therefore there was no pre-movement activity related to upcoming direction.

It is not clear from these results which brain region could be responsible for the initial response of moving to the directional port since all 5 brain regions begin at nearly the same time. The DMS and DLS do reach significance as a population, 100ms before the OFC, but it is not clear that this result is compelling enough to argue that the striatum leads the OFC. It is interesting that there is not a clear leader in the control of actions and movements between the OFC and striatum, so it may instead indicate that both regions are receiving inputs from the same brain structure about movement and action selection.

Action Encoding: Summary and Conclusions

No other study has directly compared directional selectivity, absent of reward specific movements, in all 5 of these brain regions, at the same time. It was not surprising that areas of the striatum encoded movement direction, since the Basal Ganglia has long been known for its role in action selection. However, finding simultaneous directional selectivity in both the DMS and DLS, with only the DMS showing a contralateral bias is a more novel finding. Since the DMS and DLS can be segregated based on their down stream connections, it has long been known that the DMS, in particular, would have a bias for the contralateral movements. This has been demonstrated in the past with lesion studies and recording studies. However, previous work in the DLS has found contradictory results, with some studies positing a contralateral bias in the DLS as well. Previous work in our lab with electrophysiology recordings did NOT find a contralateral bias in single cell units within the DLS, while lesion studies did find a contralateral deficit (Gage et al., 2010).

By directly comparing the two regions simultaneously, in a task that isolated movements alone, we are able to show conclusively that the DLS does care about general movement, but is doing so without a strong bias for a particular direction, whereas the DMS, as well as the nucleus accumbens Core and Shell are biased in their movement direction selectivity. Additionally, this directional encoding is not specific to the actual location of the nosepokes (spatial representation) rather it is movement specific to the egocentric location of the turn.

Not only is the striatum crucial in encoding action, the OFC had the highest proportion of cells to differentiate contralateral and ipsilateral direction. Although previous rodent work has shown direction-specific firing in the OFC, this result is not

typical in primate studies. Previous rodent work has also not always disentangled movements from spatial locations and reward-related encoding. This task was able to isolate action-specific encoding and decisively show that the OFC is involved in movement specific directional selectivity.

Even though previous research has robustly found action-related activity in areas of the basal ganglia, no study has comprehensively examined all areas of the striatum in the same task, in the same rat in a reward-related task that differentiates specific movements from reward. In this task, when examining all behavioral events, we only find direction-selective encoding during movement. All 5 regions show a high proportion of cells encoding directional selectivity- so these are cells that are not only significantly active during the movement-related events, but are also showing turn-specific selectivity. It is therefore surprising that such a large proportion of cells from all regions have this directional selectivity.

In summary, the most surprising results found were the strong encoding of directional selectivity ONLY, seen in both regions of the accumbens and the OFC. Previous research in rats have rarely disassociated specific aspects of movement vs. reward in the nucleus accumbens to be a particularly strong candidate for encoding direction, only, and so this is one of the only single-unit studies showing very specific action encoding in BOTH the OFC and NAC.

Striatal Specialization for Cue

Despite the homogenous encoding of Action seen in all 5 brain regions, differences between regions emerged early in the task in response to the Tone stimulus. It was the original prediction that the Nucleus Accumbens Core would respond most to the Tone, along with the DLS. When analyzing neural firing rates during the time of the tone, the analysis focused on the *difference* in firing between the two tones rather than examine the total activity pattern at the tone compared to baseline. The goal of this analysis was to see how the brain regions differentiated between stimuli. At the time of the Tone the DLS, Core and Shell showed a discernible response to a specific cue.

The current behavioral task focused on examining the difference in firing rates between the two tones. The tones played at an unpredictable time, and the location of the rat within the recording chamber was flexible, so tone presentation did not rely on any behavior from the rat. Additionally, presentation of the tone stimulus was only indicative of the reward the rat could work towards- it did not have any information about the expected motor movement. When the rat heard the tone, he was expected to find a lit center nose port, and then begin a motor movement on the way to potentially collecting a reward.

Dorsal Striatum: The Dorsal Lateral striatum reached a significant proportion of encoding the Cue at the time of tone presentation, reaching a maximum encoding of 10% for 600ms (Figure 9A). The Dorsal Medial Striatum did not reach significance for the Cued reward, at any point when the tone was playing. Previous work in the Berke lab has not found a robust signal for a cue in either maze tasks (Berke et al., 2009) or in a simple

reaction time task (Gage et al., 2010), so finding any cue-responsive encoding was somewhat surprising.

Despite the fact that the Berke lab has not seen a role for the DLS in cue responsive encoding, there are many reports of Cue encoding in single unit recordings in the primate striatum (Apicella et al., 1992; Vicente et al., 2012; Yamada et al., 2004); Hori et al., 2009). These results are not always consistent when studying single units in the rodent dorsal striatum. Most studies find very few, if any, cue-responsive neurons (Carelli et al., 1997; Jog, 1999; Berke et al., 2009; Kubota et al., 2009; Root et al., 2010; Barnes et al., 2005; Barnes et al., 2011). The only previous results in rodent electrophysiology studies with cue responsive neurons, when comparing the DMS and DLS is a study looking at alcohol-related cues, and finding only DMS cue encoding (Fanelli et al., 2013), or cue encoding in both DMS and DLS (Stalnaker et al., 2010).

Evidence from rodent lesion studies indicates that the DLS may be necessary for executing stimulus-response associations (Atallah et al., 2007; Balleine et al., 2007; Yin et al., 2006). However, recent electrophysiology studies in rats and mice have found conflicting results. When recording from the DLS in mice, Kubota (Kubota et al., 2009) found very few neurons that differentiated between the two auditory cues that instructed the mice where to go. They did however find an increase in only the DLS at the start and end of the task, even if the neurons did not differentiate between the two stimuli. Similarly, Barnes and Job all found very few cue related neurons in the DLS, occasionally finding none at all (Barnes et al., 2005; Barnes et al., 2011). In a task similar to the one used in this study, where the cue that instructs where to go is largely separated in time from the actual choice, Barnes (2011) found that only 6% of neurons

differentiated between the two tones at the beginning of the task. In the current study, 10% of DLS neurons responded based on the tone that played, meaning that cue-responsive neurons were not the overwhelming majority of the responses seen in the DLS.

The difficulty in comparing past results with those seen in this behavioral task can be partially explained by both the task and the analysis used. In a majority of previous research, both in the primate and rodent literature, there is often only one cue used. Most studies utilize probabilities of reward, rather than 2 separately valued rewards. A study that did utilize two separately valued rewards was Stalnaker (2010), but they only manipulated the amount or the delay to reward, whereas the actual identity of the reward stayed the same. This study did find single units that differentiated between the cues in the DMS and DLS (Stalnaker et al., 2010).

The behavioral task used in this project appears to be quite unique, then, and the finding that the DLS differentiates between two different tone-reward pairings is a novel finding. The fact that there were two different rewards may also explain why there was distinct firing for the two tones, a result that has not been seen in the Berke lab before (Berke et al., 2009; Gage et al., 2010). In these studies the tones indicated a different movement response, but the rat was still working for one specific type of reward- a sucrose pellet. In this case, the idea of stimulus-response encoding was seen to be the specific action that was taken. However in the current behavioral task, there may have been stimulus-response encoding that actually signaled the response as a differentiation in the type of reward.

Work in the rodent utilizing lesions has hypothesized a role for the DLS over the

DMS in pairing cues with responses. In a series of studies by Featherstone et al, they systematically lesioned either the DMS or DLS in instrumental learning and found that the DLS was required for stimulus-response learning (Featherstone and McDonald, 2004a, b; McDonald and Hong, 2004). Lesions of the DLS impaired the rat in discriminating between a rewarded and unrewarded stimulus, but it was only due to the discrimination between the CS+ and CS- and not a motor deficit.

It follows, then, that the DLS in the current behavioral task is responding to the Cue because it represents a pairing between a specific tone, and the response required to get that reward, that is, which port to go to. The fact that the DLS does not show a significant bias towards the higher or lower valued tone, based on the reward it is paired with, or to a specific tone-stimulus, is further proof that the pairing is due to the response selection required, rather than specific tone qualities, or value associated with the tone.

Nucleus Accumbens: When examining the two subregions of the Nucleus Accumbens, the Shell began differentiating between the two Cues at the onset of tone presentation. A significant proportion of cells within the Shell encoded 'Cue', and maintained encoding for 800ms, but the region only reached a maximum proportion of (11%). The Core had a much smaller window of significance, reaching a 12% maximum for only 500ms.

Examination of the precise nature of encoding at the time of the tone revealed that the Shell was responding by increased firing for the Water tone, regardless of internal state (Figure 9C). The Core did not reach significance for either a specific stimulus or a specific value parameter- meaning there was equal firing for both food and water tones, distributed among hunger and thirsty days.

Tone encoding in the Nucleus Accumbens has been seen in many previous studies (Roesch et al., 2009; Nicola et al., 2004b; Setlow et al., 2003; Taha et al., 2007). These results demonstrate a much stronger role for the Shell in responding to the tone, compared to the Core. This is in contrast to a great deal of previous research examining the specific roles of the Core and Shell, many of which have only found cue responding in the Core. Ambroggi (Ambroggi et al., 2011) found that iNActivation of the Core, but not the Shell, caused a decrease in responding to reward predictive cues. Additionally, when recording from the Core vs. Shell, they found more frequent, and larger magnitude firing rate responses to the reward predictive cue in the Core vs. the Shell. Both of these results indicate the importance of the Core in cue-related responding, which contradicts what was found in the current study. However, their study used a discriminative stimulus task that had either reward vs. no reward predictive cues, unlike the current behavioral task, which had slightly different valued reward predictive cues, with no 'unrewarded' cue.

Evidence that the Core is more important for cue-related behavior shows that inactivation of the Core, but not Shell, disrupts reinstatement of drug and food seeking behavior (Fuchs et al., 2004; Floresco et al., 2008). Cocaine cues cause an increase in dopamine in the core but not shell (Aragona et al., 2009). Other voltammetry studies have found that cues predictive of sucrose increase dopamine release in the Core (Roitman et al., 2004; Day et al., 2007; Jones et al., 2010; Cacciapaglia et al., 2011).

Lesion studies have also indicated the Core over the Shell in cue-guided behavior (Di Ciano and Everitt, 2001; Floresco et al., 2008; Chaudhri et al., 2010). NAC neurons show Cue related responses that correlate with both the predictive value of the cue and

the action elicited (Nicola et al., 2004b; Taha et al., 2007; Ito and Doya, 2009; Roesch et al., 2009). These results are all contradictory to the current findings- the Shell fires more for the Cue, and does so by firing for the identity of the reward, not the value. One of the only studies to find a stronger response in the Shell, compared to Core, is Cacciapaglia (Cacciapaglia et al., 2012) who found dopamine release events in both Core and Shell in response to a cue for a sucrose reward.

This study was not looking for cells that showed a more general signal for tone-related activity. The hypothesis was testing if there were cells that specifically differentiated between the food and water tone. In other studies, such as Goldstein 2012 (Goldstein et al., 2012), there were only a small percentage of cells that were active during the cue that actually distinguished between the different trial types, which is similar to the results seen in the NAc Shell and Core in this task. The Core may have a more general cue response, which is separate from a more rare response that actually differentiates between different tones.

Other studies that show slightly more subtle changes in firing for the Cue show that NAc neurons that respond to the Cue are firing based on the cost of the trial, as it relates to the delay before getting a reward of the same magnitude (Day et al., 2010). In this population, neurons exhibited a larger magnitude excitation on low-cost trials compared to high cost trials. This last part of their results is in contrast to the current results, however, since the Shell cells in this task fired for identity. A caveat to their results is that action was directly related to the cue, so these cells could actually have been distinguishing between upcoming movements, rather than the reward predictive information.

Finding more firing in the NAC Shell compared to the Core is a novel result, and makes an important contribution to the understanding of how different subregions contribute to encoding a decision-making task. This behavioral task is unique compared to most other studies that examine single unit NAC firing, because it specifically compares two different types of reward. Past results that find firing in the NAC Core and not the Shell rarely examine two separate rewards. This study adds to that literature and demonstrates a role for the Shell in distinguishing separate Cue identities, in a non-value based manner.

Orbitofrontal Cortex: However surprising it may have been to find encoding for the Cue in the Shell rather than the Core, it was equally surprising to NOT find encoding in the OFC. Primate studies from practically the beginning of electrophysiology have found that the OFC responds to salient cues that either relate to, or predict reward. Results like these are some of the most consistent findings in the primate OFC literature. Primate electrophysiology studies dating back to at least 1983 have found single unit examples that respond to different stimuli that predict rewards (Thorpe et al., 1983). Using both visual and olfactory paradigms through the years, research have consistently found neurons in the primate OFC that respond to reward-predictive cues (Rolls et al., 1996; Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006; Hassani et al., 2001; Simmons et al., 2007; Kobayashi et al., 2010). Most studies show higher firing for cues that predict a more valuable reward (Tremblay and Schultz, 1999; Roesch and Olson, 2004; Padoa-Schioppa and Assad, 2006; Roesch and Olson, 2007; Simmons et al., 2007; Bouret and Richmond, 2010; Kennerley et al., 2011; Padoa-Schioppa and Cai, 2011).

In the current study, the OFC never reached significance for the Cue at the time of the tone. Besides just the difference in species, one explanation for why there is no cue-related firing in the OFC could be due to the difference in behavioral tasks. In most primate studies, they are rarely comparing cues that predict different rewards. Additionally they rarely use a task where the monkey gets a lot of trials wrong. The studies examine only correct trials and use a behavioral task that is predictive and does not utilize a choice, using highly trained monkeys where behavior is 90% or better.

Anatomical differences between primates and rodents could also explain some of the difference in findings. The areas recorded from in this study may not line up with the precise areas seen in the primate (Wallis, 2012). In this study, the recording locations are from both medial and lateral OFC. In primate research, lateral OFC is often shown to be critical for updating the value of objects during satiation studies, while medial OFC is often crucial in studies that test the inhibition of responding during extinction (Rudebeck and Murray, 2011a; Rudebeck and Murray, 2011b). Another study comparing medial and lateral OFC in primates show that the lateral OFC is used in reward-credit assignment, while the medial OFC is necessary for reward-guided decision making (Noonan et al., 2010). In rats, lesion studies beginning to dissociate the lateral and medial OFC found that lesions to medial OFC made rats less impulsive while lesions to lateral OFC made rats more impulsive (Mar et al., 2011).

Few single unit studies have compared medial and lateral OFC in the rodent. A recent study by Burton (Burton et al., 2013) directly recorded from medial OFC in a task that used odors to predict either short or long delayed rewards. They found very few neurons that differentiated at the time of the odor cue (only 16%, or 41 neurons),

compared to baseline, and found only 13 cells that differentiated between the different odor cues. The cells that did respond to the cue in the medial OFC fired more during odors that predicted low-value outcomes. These results give insight into why the current study did not find a significant proportion of cells encoding the Cue at the time of the tone in OFC- previous work that actually tries to disentangle specific cues show very few neurons that do so in the OFC.

Other differences between the current task and previous rodent electrophysiology recordings studies are what the actual Cue stimulus was. Nearly all rodent studies that have recorded from areas of the OFC used odor cues to predict reward (Feierstein et al., 2006; Furuyashiki et al., 2008; Young and Shapiro, 2011; Roesch et al., 2006; van Duuren et al., 2007; van Duuren et al., 2008; van Duuren et al., 2009; Kepecs et al., 2008). These studies had a cue that simply predicted a reward, with no value related information, or only used correct trials. The Feierstein study had an odor cue that indicated both direction and reward and in this task, 14.5% of cells fired discriminately between the two odors. In contrast, the cue used in the current study did not discriminate movement direction. In the Furuyashiki paper, there were 4 odors, so action and odor identity could be distinguished from each other. During cue sampling, both response and outcome selective neurons were found. Odor sampling occurred 500-1000ms after initial nose poke, and rats were required to respond within 5 seconds. The rat then had to hold for another 500-1500ms before receiving reward. The odor sampling port and reward delivery port were on the same wall, so movements were very quick, despite any delay imposed. Similar to the primate studies, the time between a cue stimulus sampling and actual reward receipt was less than 5 seconds, so anticipation of reward would be hard to

separate out from the cue response, especially since rewards were delivered to the same well.

In contrast, the behavioral task used in the current study requires the rat to make an extended movement in between the time he hears the cue and receives the reward. The timing difference between this task and other rodent and primate studies may be why there is a difference in results for OFC cue-related firing. One possibility is that previous results misinterpreted a reward anticipation signal as cue-related response instead. The other possibility is that the OFC is not engaged in a task that has an extended time delay between Cue presentation and actual reward receipt. Further analysis examining trial lengths would be an exciting extension of the current results.

Evidence that suggests previous research may have had an overlap between cue response and reward anticipation can be seen in the current analysis, which shows that the OFC does fire for reward anticipation beginning 2 seconds before goal port entry (see Goal Chapter, Figure 12A). Relative to that time frame, previous results that have a short duration between cue presentation and reward may simply be washed. Additionally, since the rats in the current behavioral task are allowed to choose the reward port, and often get the trial incorrect, it is more than likely that the Cue that plays is not very predictive of the actual reward.

Previous studies done in the OFC from the van Duuren group, which looked at reward probability during odor sampling found that 7% of neurons changed firing rate during odor sampling. Due to the nature of their task, they didn't examine reward probabilities associated with the rewards, since they couldn't dissociate odor identity from probability (van Duuren et al., 2008; van Duuren et al., 2009). They did find that on

a population level, as well as on a single unit level, the reward probability was encoded during the movement, waiting and reward periods. This study is one of the few rodent studies that did not find a very large number of cells in the OFC that encode the cue at the presentation of the stimulus. They impose longer delays (1.5s once the rat enters the reward trough) before delivering reward, so in a similar manner to my study, it may mean that if there is a longer waiting period between hearing the cue and receiving the reward, fewer OFC neurons respond to the Cue. Additionally, they had a reward probability task, so reward wasn't guaranteed after odor sampling, unlike some of the other rodent single unit studies, which may also explain why their study, like mine, doesn't find a large fraction of OFC responsive neurons during Cue presentation.

Overall Cue Discussion:

The difference in behavioral tasks most likely explains a lot of the results either seen or not seen in the current results- this task is not probabilistic, but it is also not strictly Pavlovian. Most studies in both primate and rodent played a reward predictive cue that always meant a reward was coming- there was often no choice involved. In other groups of studies that have found cue related activity, it was for tones that compared reward vs. unrewarded stimuli- in this case it means there is no choice, the rat/primate will either receive a reward or he will not. The difference with the current behavioral task is that even though the rat heard the cue, and technically understood what the correct response was, it did not make receiving that reward a forgone conclusion. The rat could still enter the incorrect reward port, and not receive any reward. However, getting the trial wrong is still different than hearing a cue that indicates, ahead of time, that the trail

is unrewarded. The current task still does not make it a probabilistic task, because even if the rat still made a choice about which goal to enter, he technically had the information available to get every trial correct.

A task that is similar to this task uses a t-maze to impose a choice for the rat, and uses a tone to indicate the correct direction. In these studies, there are very few reports of actual discrimination between two tones (Thorn et al., 2010; Barnes et al., 2005; Barnes et al., 2011).

Another difference between the current task and most other simple instrumental tasks that have found cue encoding is the length of time between when the tone actually plays to when the rat is making a decision about which reward port to enter. The average time it takes to complete a trial ranges from 6.5 seconds to almost 14 seconds, depending on the individual rat. This is much longer than most nosepoke, saccade and lever reaching tasks.

An even more compelling reason that the current task is different than other tasks is the fact that not only were incorrect trials utilized, but also incorrect trials were repeated. This means that approximately 10-25% of trials were repeated from the previous trial. The rat had information about the upcoming trial before the tone actually played. Over time, this may have had the effect of making the Cue a less valuable stimulus. Pair that with the fact that the rat was wrong on 10-25% of trials means that the predictability of the cue is even lower. When the 3-factor regression was analyzed using only trials that were previously incorrect, the response to the Cue no longer reached significance in any brain region (Figure 10A). When analyzing Cue response in only the non-repeat trials (so Previously Correct Trials) the same proportion of cells becomes

significant- there is no alteration in encoding seen across the brain regions (Figure 10B). The Shell and DLS are only responding to the Cue when there is no previous information about the current trial.

On Sated and Restricted days, for Cued only, the DLS reached a higher proportion than the Shell, and encoded the Cue for longer (Figure 8B). The Core also became significant for a longer period of time. However, there was no clear difference in the index for food vs. water reward. This demonstrates that the firing at the time of the tone was for some stimulus identity, or possibly signaling a salient event. The response is not dependent solely on value, since the encoding of the tone still happens on sated and restricted days, when there is little difference in value between the two tones.

Specialization for Action-Goal in the Dorsal Medial Striatum

Although Action encoding was universal across brain regions, the integration of Action and the upcoming goal was only seen in the Dorsal Medial Striatum, demonstrating another instance of regional specialization. Previous research studying the integration of action with reward has found a signal in the dorsal medial striatum for 'Action-Value' (Samejima et al., 2005; (Lau and Glimcher, 2007; Lau and Glimcher, 2008; Pasquereau et al., 2007; Seo et al., 2012; Kim et al., 2009; Kimchi and Laubach, 2009b). In a majority of these studies, animals were given a probabilistic choice task, with no actual correct answer. There was no cue that told the rat which goal port, or action was correct so the rat had only had a certain probability of finding a reward. Different movements were assigned different probabilities for reward, and the typical task design included blocks of various reward probabilities, so that specific action-value terms, calculated using equations from reinforcement literature, could subsequently be examined.

The current behavioral task was not designed to specifically test for 'action-value'. Instead, this task was designed to disentangle movements from rewards, in order to test how those factors are encoded without being linked to one another. Additionally, this task was set up so that the rat was always informed as to what the correct answer was. The difficulty came from the fact that the rat's desires often interfered with the correct choice. This led to behavioral errors in the 10-30% range, depending on the individual rat and trial type (see Figure 3E).

Interestingly, despite the fact that actions were not specifically tied to rewards, or reward probabilities, there was still an integration of action and choice, based on which goal port the rat went to. This result is similar to an action-selection signal, or an action-choice signal, rather than an action-value signal, since it integrates both the direction and goal port the rat actually went to (rather than the option the rat faced). This also happened at a time that was before actual reward port entry, so the integration of action and upcoming goal port could conceivably be the actual action-selection mechanisms- or choice-like mechanism. This was seen mostly in the DMS, and very briefly in the OFC.

The fact that the DMS signal was earliest and strongest could mean that the actual decision was being made in the DMS about where to go. The fact that the cells fired most on non-preferred trials could mean that the DMS is only involved on trials that require the most effort. Other studies attempting to distinguish the first instance of action-selection between the OFC and Striatum found that the OFC fired first (Simmons et al., 2007). However, in that study, they were comparing different animals and studies done at different times. The current study is the first conclusive finding that directly compares the OFC and DMS in a task that can determine where an action-goal signal originates.

Findings of action-goal in the DMS are not new, however, it was unexpected to find this signal *only* in the dorsal medial striatum and OFC. Other rodent studies in particular have found an action-value, or action-outcome signal in the Nucleus Accumbens (Roesch et al., 2009), and due to its connections, the NAc is often thought of as the limbic-motor interface (Mogenson et al., 1980). In contrast, follow up studies that specifically disentangled actions and goals, like the current study, did NOT find action-goal related firing in the Ventral Striatum (Goldstein et al., 2012; Ito and Doya, 2009).

Other rodent studies have found an action-goal signal in the DLS (Stalnaker et al., 2010). In this study they used a free choice and forced choice task to separate actions and responses, as well as the cues that predict them, and found that both the DMS and DLS represented action + goal contingencies in the time before receiving reward. In their analysis, they looked at response selective neurons, and then examined the differences in firing rate during different predicted rewards. This analysis did not actually use statistics- they just showed the average firing rates of different types of neurons, collapsed across trials.

The current studies unique ability to disentangle all aspects of reward related information, allow a more focused analysis on the important aspects studied in decision-making: Cue, Action and Goal. Using this task, it is clear that the DMS has the strongest role in combining actions with upcoming Goal. These are not cells that only react to the Cue, nor are they cells that only respond to direction- they specifically integrate Action and Goal, regardless of what the upcoming outcome will be- this signal is very specific to the movements and goal port.

To further understand how the DMS was encoding Action-Goal, we examined the firing rates of the individual cells and found that nearly all cells that reached significance in the 3-factor regression were firing most for the contralateral movement on non-preferred trials. This signal is rather puzzling, because there are very few studies showing any brain region that consistently fires more for a non-preferred reward. There is some evidence though, that this signal may be necessary to help guide an animal through less preferable trials onto a more preferable trial (Minamimoto et al., 2005; Minamimoto et al., 2009). In these studies, the authors propose that the DMS is involved in 'rebiasing' an

action that is required in order for the rat to correctly go to a smaller reward, since the brain is usually biased to collect higher rewards. Corresponding to this, the results seen in the current study show that a majority of the neurons with an action-goal significance had the highest firing rate on Correct, Contralateral, Non-preferred trials (Figure 11C). It appears that the rat is putting in an effort to go to the non-preferred reward port to get it correct. This may have to do with the nature of the task because incorrect trials are then repeated. Possibly, the rat is putting in more effort to get the non-preferred reward correct so that he can then get a preferred reward the next time, rather than repeating the non-preferred trial, similar to the theory proposed by Minamimoto et al.

It could be that the DMS is rebiasing the rat into picking the non-preferred reward. Another interpretation is that the DMS is exerting extra control, or effort, on the non-preferred trials due to the fact that if incorrect, the current trial will be repeated. That means that an even longer delay will occur, and the rat will not receive the preferred reward for an even longer period of time, making a mistake very costly, in terms of time. In order to actually get to a preferred trial, then, the rat must exert greater effort on those non-preferred trials, and this is encoded in the DMS. The DMS is exerting cognitive control on trials that are against the Pavlovian response of simply heading to the more preferred reward port.

Instead of firing for a combination of action + goal, the neurons could have fired for an integration of action + cue. This would be similar to the action-value like encoding seen in Samejima (Samejima et al., 2005) and Lau and Glimcher (Lau and Glimcher, 2007) because it would have been an alternative option of task-related factors, other than simply where the rat chose to go. Individual brain regions could have encoded the Cue in

order to maintain a signal for the correct reward during each trial. Since most rats were incorrect on 10-25% of his choices (Figure 3E), there were enough trials to actually compare what was cued to what the rat chose (goal). In terms of the RL literature, this is similar to having options, in a reward probability task, even though the task was not specifically designed this way.

The neural firing rate patterns followed the rats' behavior, and not the Cue. The neurons that fired during this time frame did not encode the cued tone- there were no brain regions that fired differently for the cued tones while also firing for action (Figure 11E). The result is somewhat surprising, given that recent studies in the rodent and primate literature have demonstrated that various choices are encoded in the striatum (Samejima et al., 2005; Lau and Glimcher, 2007). The fact that this task was not a probabilistic choice task could explain why there was no integration of the Cued reward prior to entry into the goal port. Another explanation could be that the rats always thought they were going to be rewarded, which is why they always encoded the upcoming port choice, rather than the actual Cue that played. The final explanation may be that there was too much time between when the tone played and when the rat made the final movement to the reward port. The average trial time was between 6-11 seconds, depending on the rat, so this may have exceeded the working memory needed to actually encode the Cue at the time of the decision.

Despite the fact that the Nucleus Accumbens is often referred to as the 'limbic-motor interface' (Mogenson et al., 1980), there was no encoding of action-goal, action-cue, or action-outcome related variables in either area of the Nucleus Accumbens. Cells

that reached significance only did so for the main effect of either term- Action or Goal- with no overlap between the two.

Anatomical studies (Heimer et al., 1991; Voorn et al., 2004) as well as lesion studies (Cardinal et al., 2002b) have suggested that the NAc participates in integrating information about the value of expected rewards with the motor behaviors needed to get the reward. Other work in RL related fields, however, suggest that the ventral striatum has a role in encoding value as a critic, and therefore would relay that information downstream, rather than acting in an integrative manner (Takahashi et al., 2008; Padoa-Schioppa, 2011).

Few studies have actually examined reward related information in the ventral striatum without having the motor output linked directly to rewards. Studies either examine probabilistic rewards, or change the instrumental response. Studies that presume the NAC acts to integrate motor and reward related information utilized tasks that gave action + reward specific directions (Carelli and Deadwyler, 1994; Bowman et al., 1996; Setlow et al., 2003; Janak et al., 2004; Nicola et al., 2004b; Shidara and Richmond, 2004; Taha and Fields, 2006; Simmons et al., 2007; Ito and Doya, 2009; Kim et al., 2009; van der Meer and Redish, 2009a; van der Meer et al., 2010; Day et al., 2011).

Utilizing a task that did separate the cue signaling reward from a directional response, in a similar manner as the current task, Goldstein (2012) found NAc neurons that encoded value independent of the motor behavior. The task differences were the fact that the cue was an odor presentation, and the rewards were just different magnitudes of the same liquid reward (3 drops vs. 1 drop). Their task was also quite fast, possibly lasting 1-2 seconds at most. Even if the task was fast, they found that rats were slower on

the more preferred trials, again indicating that separate populations of neurons in the NAc may encode value, independent of a motor, or motivational response.

Ito and Doya (Ito and Doya, 2009) also recorded from the NAc and found a small fraction of neurons encoding action-value, but like a lot of other studies, they used a probabilistic task that relied on actions that specifically led to reward.

The direct comparison of the current task with other previous research is not straightforward. This task specifically manipulated independent aspects of motor and reward related response, and also utilized behavioral errors to examine choice and value. Because of these manipulations, it makes a strong case against either action-goal or action-value information in the nucleus accumbens. If the NAc held the opposing options, while integrating action, then we would have expected to see encoding of some mixture of action, cue, and the interaction between the two.

Only the DMS integrated the movements/direction the rat went with the upcoming goal choice. It did so in a manner that suggests it was actually an action *selection* mechanism, rather than integrating information specific to the potential outcomes- since the signal only showed up by examining where the rat chose to go. It could be that if the Nucleus Accumbens is involved in representing action-goal, it can only do so in a task that links very specific actions with their outcome- that only one type of action can lead to one specific reward. The DMS on the other hand, can still integrate action and goal when they are not explicitly linked. This is evidence that the DMS is responsible for the actual action selection process.

Summary and Conclusion: Rebias for Non-Preferred Reward in the DMS

Actual evidence for action-selection in the DMS has rarely been seen in rodent studies. Kim et al. only found a brief indication of the upcoming choice in a free-choice maze task, 50ms before the choice was manifested (Kim et al., 2009). Stalnaker et al found action-goal encoding in both the DMS and DLS (Stalnaker et al., 2010). Seo, Lee and Averbeck found action-value representation in the striatum, but found a strong action selection activity only in the lateral prefrontal cortex (Seo et al., 2012).

Previous research has found an action-value signal in the DMS. The signal found in this study in the DMS is for action + goal integration. This signal is specific to the contralateral movements paired with the non-preferred reward. The signal seen here is evidence of the DMS exerting cognitive control in an action-outcome encoding scheme that is utilizing a model-based framework to specifically re-bias, or to direct the rat to choose the Non-Preferred reward port correctly. The DMS is required to exert more effort on the less preferred trials, in order to not make a costly mistake and go to the 'Pavlovian' port- the port that on that day delivers the more preferred reward. Additional support that the DMS truly is being utilized for action selection is the fact that this comes before entry into the reward port, and it is only for the goal port the rat chooses- not for the actual outcome, or the cued reward.

Universal Encoding of Upcoming Goal Port Choice

Despite specific striatal subregion differences in Cue and Action-Goal integration, the encoding of which reward port the rat was about to enter was another signal shared across regions. As the rat finished the movements to the adjacent nose ports, the next event in the task was to enter a reward port, and potentially receive a reward. In this task, the rat was cued as to what the correct port was to enter, but since there were two separable reward ports, he actually had a choice prior to reward port entry, and could decide not to enter the cued (correct) port. An incorrect choice causes the houselight to come on, and a lengthy time out until the rat is allowed to begin the next trial. Additionally, if incorrect, the rat must repeat the trial until he is correct, which creates an additional delay. This was crucial to the training and performance of the task in the rats, because their behavior was strongly influenced by internal state, so often on trials when the non-preferred reward was cued, the rats chose the preferred port, incorrectly. On a very small minority of trials (<11 trials, on average), the rat incorrectly went to the Non-preferred reward port when the preferred port was cued.

When the rat chose the incorrect reward port, the Cue that played differed from the Goal chosen. Since the rats consistently made mistakes due to internal state, rather than simply not understanding the task, behavior was stable, which allowed us to compare the how a decision was made as the rat approached the reward. The potential neural correlates of differential activity could be accounted for by the Cue that played- the rat could be remembering the tone, and continue to fire for that tone, regardless of if he entered the correct port. Alternatively, the firing rates could instead be a neural correlate for Goal- that is, the actual reward port the rat entered, regardless of which tone

played, and therefore regardless of if he would get the actual reward. Finally, he could encode the interaction between Cue and Goal, which serves as the signal for 'Outcome'- if he went to the Cued Goal port, he was correct, which causes an interaction between 'Cue' and 'Goal' in the multiple regression analysis, while going to the wrong port still has an interaction effect, it is just in the opposite direction.

Surprisingly, the only neural signal that reached significance before the rat entered the reward port was for Goal only. Starting long before the rat even entered the reward port, all 5 brain regions began firing more for the port the rat would enter, rather than fire for the port that was cued. The DLS and DMS began encoding the upcoming chosen reward earlier than the other brain regions, beginning almost 2 seconds before reward port entry. The Core, Shell and OFC encoded Goal more than 1 second before the rat got to the reward port, indicating that all 5 regions carried information about the reward port the rat intended to enter.

The original hypothesis was that nearly all brain regions would encode the upcoming anticipated reward, prior to reward port entry. This encoding would most likely reflect the 'Cue', since that was the correct reward, and it was signaled ahead of time. Depending on the task, previous research has found signals relating to upcoming reward probability, upcoming value, or upcoming choice, in the rodent OFC, rodent dorsal striatum, as well as the rodent ventral striatum (van Duuren et al., 2007; van Duuren et al., 2009; Sul et al., 2010; van der Meer and Redish, 2010; van der Meer et al., 2010; McDannald et al., 2011). This has not been a universal finding in the dorsal striatum or the Nucleus Accumbens, with some studies not finding a choice signal prior to the behavioral manifestation of the decision (Thorn et al., 2010; Kim et al., 2007).

Additionally, no single unit studies have been conducted in such a wide range of cortical and basal ganglia areas, so we could only make a guess that the OFC would signal this choice earliest, due to a previous study comparing striatum and OFC in primates (Simmons et al., 2007) as well as an early signal in a free choice task in the rodent arganular cortex, which projects to the rodent OFC (Sul et al., 2011).

Much to our surprise, all 5 brain regions significantly encoded 'Goal', rather than the other options of 'Cue' or 'Outcome'. In fact, no region even became significant for only the 'Cue' either before or after reward port entry. The specific signal we found was only related to whether or not the rat entered the food port or the water port, which are physically separated on the opposite wall of the behavioral chamber from where the nose poke ports are located. The reward port locations did not change for any rat throughout training and testing and therefore have a very strong association to reward delivery.

One interpretation of this signal could be for the location of the ports, similar to 'place cells', since they are in physically different locations. This interpretation is not likely due to the fact that the differentiation between ports is seen as the rat leaves the last nose poke, which is typically 1-2 seconds before entering the reward port. If it were a specific place signal, it would only differentiate at the time the rat enters that physical space. It is also most likely not a 'turn' signal, or a movement related signal, because the rat makes different contralateral and ipsilateral movements, regardless of his starting location. The way that the reward ports are physically separated from the nose poke ports also means that the rat has completed a contra or ipsi rotation before heading directly to the port, meaning that the turns used here are categorically different than the contra/ipsi

nosepokes made at the beginning of the task.

Instead, this signal most likely relates to some specific aspect of anticipation for the upcoming reward, based on that reward's identity. The fact that there was such a strong encoding of the upcoming goal choice 2 seconds before the decision was revealed that was a specific identity signal, rather than a general 'reward' signal is a novel finding. It is also a unique finding because this signal was not related to the Cue or the Action prior to entering the goal port, yet was significant nearly simultaneously in all 5 brain regions. Most previous research on rodents has not directly compared two different stimuli when examining reward-related firing (Wallis, 2012). Instead, other behavioral tasks either utilize a cue that indicates reward vs. no-reward (Taha et al., 2007; Nicola et al., 2004a), or the task involves some probability of reward (Sul et al., 2010; Sul et al., 2011; Ito and Doya, 2009; Setlow et al., 2003). In those tasks, neuronal firing rates are then correlated to either reward anticipation, or reward probability. Additionally, most other previous research has not specifically isolated expected outcome from the behavioral response, or movement (Day et al., 2006; Nicola et al., 2004a; Ambroggi et al., 2008; Feierstein et al., 2006; Roesch et al., 2006; Young and Shapiro, 2011). The current behavioral task finds specific neural encoding of the upcoming goal, irrespective of the action it takes to get there, and this signal is not related to a general signal for reward anticipation- it is specific to the identity of the reward.

Finding coding for distinct reward identities is not a novel result. Previous work in both rodents and primates has shown that when animals are working for different categories of reward (i.e. juice vs. cocaine), areas of the striatum respond with different subpopulations of cells (Carelli et al., 2000; Carelli and Ijames, 2001; Cameron and

Carelli, 2012; Bowman et al., 1996). In these studies, the theory has been that the brain actually responds differently to 'naturally' reinforcing stimuli compared to drugs of abuse (Carelli et al., 2000; Bowman et al., 1996). Only a handful of studies have compared a wider variety of rewards that do not include drugs of abuse, and found nearly similar numbers of identity-only encoding cells (Carelli et al., 2000; Hassani et al., 2001). Rather than differentiating between natural rewards vs. drug rewards, it appears that the striatum and OFC have significant proportions of cells that encode identity, regardless of value.

Carelli (2000) shows similar results to the current study, finding that 32% of nucleus accumbens neurons differentiate between food and water rewards. In that study they were focused on how 'natural' rewards were different than cocaine rewards, so they did not go into great detail about how, specifically, the food and water rewards were encoded. A similar study performed in primate compared cocaine reward to a juice reward, and recorded neurons in the ventral and dorsal striatum (Opris et al., 2009). Like the Carelli study, they found differential encoding depending on reward identity in 20-30% of task responsive cells in the DS and VS. In a study done in the primate striatum (Hassani et al., 2001), 5 different types of reward were used and they found that 1/3 of striatal neurons recorded had different levels of task-related activity depending on which reward was predicted. This study did find higher firing rate in some non-preferred rewards that again show the importance of identity over value.

It is important to realize that a full 30%, give or take, of all neurons recorded from studies that have specifically examined different reward stimuli respond to the specific reward identity, rather than simply the value. A lot of previous research trying to relate specific reward activity to a reward probability or anticipation of a reward based on an

assumed 'value' of the reward may only be responding to an upcoming reward identity. The current study has manipulated value in a unique way, where internal state dictated what reward was more preferred. It was our prediction that neurons in the NAC and DMS would track the change in reward value based on internal state. As it turns out, within the neurons that clearly differentiated between food and water rewards, it was the identity, not value, that was encoded in the individual neural firing rates.

The fact that this behavioral task had the rat working for two different rewards may be why there was an upcoming goal/choice signal, whereas other single unit studies that used different goal port locations, but a single reward type did not (Thorn et al., 2010; Kim et al., 2009; Sul et al., 2010). In the Kim and Sul studies, the rats were in a free choice maze task, where different turn directions had different probabilities of reward. Unlike the current study, they found very little choice signal in either the DS or NAc before the rat's actual choice was revealed. The difference may be due to the difference in task. The rats were given no cue instruction, and their decision at the maze intersection was still not guaranteed to get them a reward, since it was a probabilistic task.

In the Thorn (2010) study, as well as Barnes (2005), they used a t-maze task, but this time there was an instruction about which direction to go, and if the rat went the correct way, he would get a reward. Like the Kim and Sul papers, Thorn and Barnes did not find a robust choice signal prior to making a decision. The Thorn study was most similar to the current study, in that there was an actual instruction about the correct direction. However, they did not find a robust decision making signal, and this may be due to the fact that there was only one type of reward. The fact that the rats were only

working for a single type of reward is the main difference between the Thorn and Kim studies and the current study, and could explain why neither of those studies found a choice signal before the rat made a decision. In the current study there is a signal before the decision is revealed, and it may be because the rat is keeping in mind the actual identity of the reward, which is only manifest if there are two competing reward identities available. If the other tasks had different reward identities rather than a single reward type, there may have been a choice signal, since that may be more relevant to what the striatum actually encodes- reward identity.

Even though the current study utilized two different rewards, it is still relevant to compare these results to studies that have looked at upcoming reward signals in a single reward paradigm. The main caveat is the fact that most previous studies used a vastly different behavioral task that was a probability task, a Pavlovian task, or an instrumental task.

Dorsal Striatum*:* Kimchi and Laubach used a go/no-go task and found around 44% of DMS neurons that were modulated around nose poke exit on Go response trials, while 33% were modulated on no-go trials. This proportion of DMS cells is slightly higher than the response rate found in the DMS in the current study. Kimchi and Laubach only used a single reward, and it was not a choice- the rat just went to the reward port to collect the reward (Kimchi and Laubach, 2009b). In their study, they also found that 28% of neurons in the DMS were modulated by the outcome of the current trial, which is very similar to the proportion of neurons that found in the DMS, after reward was received (see 'Outcome' section, this chapter).

Kim et al. tested for choice signals in a dynamic two armed bandit task, but only found signals in the dorsal striatum 200ms prior to when the choice was revealed (Kim et al., 2009). This task was different than the current task in a few ways. First of all, there was no correct answer since it was a probability task that used free choice- so there was no instructional cue. Second, there was only one type of reward possible. They state that action selection may not be conducted in the striatum, since they did not find any evidence of the choice signal more than 200ms before the choice is revealed. It could be that without prior instruction, actual action selection is not as strong or early of a signal in the striatum. However, when an aspect of either working memory, or reward identity is required, the striatum comes online at a much earlier time point, and neural signals correspond to the upcoming reward identity, rather than just a choice in turning direction. In the Kim task, the choice was a turn, but the rats may not have had to really change their behavior much more than the running motion they were currently engaged in, so a difference in firing rate may not have been detectable. In the current task, where there is a choice between actual types of rewards, the choice is revealed much earlier because an aspect about the actual outcome- the identity- was kept in mind.

In a study specifically testing the differences in DMS vs. DLS, Stalnaker et al found that 31-32% of both DMS and DLS neurons encoded information about the upcoming response (Stalnaker et al., 2010). In this task, movement and value were not separated, but they examined the firing rate differences based only on value, so the signal was more similar to an upcoming goal (or outcome) signal. Despite not having two different types of rewards, the DMS and DLS did encode information about the value of the upcoming reward prior to realizing that goal.

Nucleus Accumbens: Kim (2009) studied the Nucleus Accumbens and did not find corresponding choice signals in the NAC before the appearance of the behavior revealed (Kim et al., 2009). This is in contrast to the current results that found signals in both Core and Shell that encoded the upcoming reward port 1-2s before entry into the port. Another study examining impending goal choice was done by Kim and found that there were very few cells that encoded the impending goal choice in a forced choice t-maze task (Kim et al., 2007).

A study recording from the ventral striatum in rodents by Roesch, found that NAC neural firing rates were modulated by the value of the expected reward, prior to reward port entry (Roesch et al., 2009). This study did not disentangle left and right movements from reward value, so it is hard to interpret this as a purely reward-related signal. In a follow up study slightly more similar to mine, where they specifically disentangled actions from value, Goldstein (Goldstein et al., 2012) found that during odor presentation (a time point prior to reward entry), the NAC had higher activity during cues that predicted a higher reward. However, in this study, they did not report on the time before the rat actually receives the reward, so even though the timing was fast, there was still a delay before reward receipt, therefore it is hard to tell if the information about upcoming reward was encoded. The fact that reward value encoding was seen during cue presentation was similar to the current study, since it was absent any action-related information. However, they only used one reward type, so they were not comparing different identities of reward information, rather, they were only studying how reward magnitude modulated neural firing rates.

Ito and Doya also found a signal in the NAC prior to the rat actually getting the reward about whether or not the trial will be rewarded (Ito and Doya, 2009). Wilson and Bowman recorded from the NAC and found neurons distinguishing between different reward identities, but they were comparing rewarding vs. aversive events (Wilson and Bowman, 2004). They found that of the 30% of neurons that responded to a conditioned stimulus, 2/3 of those encoded the difference between the upcoming rewarding or aversive events.

Orbitofrontal Cortex: Single unit studies have also been conducted in the rodent OFC to determine if upcoming choice signals or reward are encoded. Sul et al. used a two-armed bandit task and did not find a signal related to upcoming choice in the OFC before the behavior was manifest (Sul et al., 2010). In a slightly different task, the van Duuren lab found that reward probability was encoded in the rodent OFC prior to reward entry (van Duuren et al., 2009). Roesch et al. used a delay task to find that firing in the OFC did predict the upcoming size of the reward, but did so differently than a delay to reward (Roesch et al., 2006). Furuyashiki also found outcome selective coding when comparing sucrose vs. water reward, beginning at cue presentation (Furuyashiki et al., 2008).

Value vs. Identity:

Interestingly, in many of the previous studies that did find some upcoming reward signal, the authors mentioned that the identity of the reward was encoded, and not necessarily the value. For example, in the DMS and DLS, Stalanaker (2010) found that in both regions, the outcome modulation did not reflect the relative value. In Ito and Doya

(2009), in the rodent NAC, they found a larger response in the no-reward predicting tone, and again, found more information about the specific trial type rather than any other combination of factors. In the OFC, Roesch (2006) found that there were populations of OFC neurons that distinguished between the large and small reward, but did not actually fire more for the larger reward.

To better understand what was encoded in the brain regions that had single units that fired for the goal port, we utilized the firing rate index to compare signals for 'Value' vs. 'Identity'. If a region, as a whole, was more interested in the value of the upcoming goal, than it would have more neurons that fired for the food reward on hungry days, and more neurons that fired for the water reward when thirsty. If, on the other hand, the signal was simply about reward identity, then there would be no bias based on internal state, and instead, a bias may show up for a specific reward identity (food or water, regardless of preference based on internal state). To our surprise, we found that the Value index did not show a bias as the rat approached the reward port. There was also no bias in the Identity index either, *prior* to reward port in. After the rat entered the port, there was then a significant bias in the Identity index for some of the brain regions. The DMS, DLS and OFC all reached a significant bias for the Food Goal Port, while the Shell and Core reached a significant bias for the Water Goal Port (Figure 12C).

This finding was not what we expected- we hypothesized that the DMS and the accumbens would differentiate encoding based on the value of the reward ports, prior to reward port entry and the value status would be dictated by internal state. Previous research in all 5 brain regions indicate that firing rates often distinguish from each other based on value (Setlow et al., 2003; Nicola et al., 2004a; Nakamura et al., 2012;

Pasquereau et al., 2007). This is not to say that this is the first study to show more firing related to specific trial type or identity, however. Ito and Doya (2009) examined rodent NAC and found that most encoding was related to information about specific trial type, rather than any type of action-value like encoding. Stalnaker (2010) also tried to differentiate between DMS and DLS in a task that independently manipulated response-outcome and stimulus response in an attempt to distinguish how each region encoded value vs. identity, and found that there was little difference between the two, finding that outcome modulation did not reflect value, but instead signaled identity. In the rodent OFC, Furuyashiki (2008) also found single cells that distinguished between upcoming rewards- sucrose vs. water- independent of the motor contingencies for the reward, and the signal began at the presentation of the cue indicating which reward the rat would receive, and signals were independent of value. Roesch (2006) also found a group of neurons in the rodent OFC that fired equally for both sizes of reward during a long delay, granted- these were the same reward, just different sizes.

In the current study, any bias that was found only showed up right at or after reward port entry- there were no regional biases prior to entry to the reward port. The biases that began after reward port entry were not responding to reward consumption because they were not differentiating between correct and incorrect trials, which would be necessary for a eating or drinking only response. This is interesting because this means that the cells responding to Goal Only after reward port in, are not responding to the actual outcome of the task. These cells were only firing for the identity of the port itself, either as a true 'port' identity, or what the port represented- food or water. However, there still are some % of cells from each region that are showing reward port

differentiation prior to actual entry into the ports, they just do not show a bias for either value or identity until the actual reward port entry.

What do these biases mean? Why is there an increase for water encoding in both regions of the Nucleus Accumbens? Previous research done in the Kelley lab and the Carelli lab show that there could be specific differences in the striatum between food and water reinforcers. Despite the fact that most results from the Carelli lab examine natural vs. drug reinforcers, and that in general, they try to lump food and water together as a natural reinforcer, a closer reading of the results are actually in line with the results seen in this study. Specifically, despite the fact that a majority of cells encode food and water reinforcers on a lever press task in a similar manner, there still are between 20-35% of task-responsive nucleus accumbens cells that did differentiate between either food and water (Carelli et al., 2000), or water and a sucrose solution (Roop et al., 2002). Further evidence for the NAC playing a specific role in water encoding comes from the fact that when previous studies compared the 'natural' water reinforcer with a drug reinforcer such as cocaine or ethanol, distinct groups of cells in the NAC only increased firing for the water reward (Robinson and Carelli, 2008; Carelli et al., 2000; Carelli, 2002). In this group of studies, they did not specifically target the Shell vs. the Core, which is similar to the fact that in the current results, both Core and Shell increase firing for the water reward. The proportion of neurons seen in this study and in past studies is also similar- right around 30%. When the Carelli group tried to differentiate between water and cocaine reinforcement in Core vs. Shell (Carelli Synapse 2006), they did not find a specific distribution between regions for cells that encoded the two reward types differently. A primate study also compared cocaine vs. juice rewards and found groups of

cells that responded differently depending on the identity of the reward (Bowman et al., 1996). These studies demonstrate that not only are there groups of cells that differentiate due to specific reward identities, but there is also evidence that some cells also show a selective increase that is specific to a water reward.

Studies that specifically address the role of the Accumbens in approaching and consuming food and water demonstrate the confusing relationship between single unit studies and regional manipulations. Research into feeding behavior has shown that iNActivation of the medial shell results in an increase in food, but not water intake, while iNActivation of the Core does not increase food intake (Basso and Kelley, 1999; Soderpalm and Berridge, 2000; Stratford and Kelley, 1997; Taha et al., 2009; Stratford and Wirtshafter, 2004; Stratford et al., 1998). However, amphetamine injection does increase responding for a water reward (Covelo 2011). The apparent increase in responding for the water reward seen in this current study by both the Core and the Shell may instead be a decrease in response to the food reward. The specific analysis that was done would not be able to differentiate between phasic excitations vs. inhibitions, so if the Shell is more responsible for feeding, but does so by actually shutting down firing, this would manifest in my data as an increase in water reward responding. Evidence from studies that have specifically examined inhibitions vs. excitations in feeding behavior have shown that different groups of NAC cells have separable firing rate patterns (Taha and Fields, 2005; Taha and Fields, 2006; Krause et al., 2010).

It is quite possible that the areas that were recorded from in this study come from two separable groups of cells in the nucleus accumbens. One group is responsible for feeding behavior, by inhibiting cell firing, while another group responds to water

rewards with a phasic increase in firing. The analysis that was done would then combining the data from these two groups of cells resulting in what looks like a water port preference, when in reality, there is a separable responses to two different ingestive behaviors- consumption of food vs. consumption of water. The important caveat to continue to remember in this assessment is that this effect is only seen in the 'Goal' cells, so they are responding to the actual port entry, and not necessarily to consuming the reward. The results would then indicate that the rat is responding to the anticipation of a food or water reward, and that this is more in line with an approach response, rather than the actual consummatory behavior.

Food Identity Encoding:

An explanation for why the DMS and DLS are firing more for entry into the food port could possibly be because, as a general rule, the Food reward maintained a higher value. When the rat was food restricted, this was obvious. However, when the rat was water restricted, he could also only eat while he had access to water. To control for this, the rat had free access to both food and water for an hour each day, as part of the internal state manipulation, but it is possible that he still carried over too much food restriction, on a day-to-day basis. This also shows up in the total trial times for some of the individual rats where they are actually faster on food chosen trials, on water restricted days (Figure 4B). Each individual rat was kept at a weight well above what a normal food restricted rat typically weighs, most rats were 500-600grams so they were not starving. It may just be that water was an immediate, pressing need on water restricted days, but as soon as that need was met, the food instantly became the more valued reward.

If this is the case, then the increase in firing for DMS, DLS and OFC could be explained in this manner, and the NAC data could also be explained as a 'dip', rather than an increase (similar to what was proposed before). However, the most obvious way to test this theory would be with other internal state manipulations. Part of our motivational protocol was to also test the rats when restricted on both food and water, as well as when sated on both food and water. Behaviorally, the rats showed a preference to the food reward- on free choice trials and on their error trials (Figure 5A and 5B). This was not as strong of an effect as was seen on food restricted days. However, the neural data in the DMS, DLS, OFC and Core are not fully explained under the umbrella 'food is always more valuable', as I will talk about next.

Sated and Restricted:

In order to examine if food really was more valued, we examined the proportion of cells that differentiated between the two rewards on sated and restricted days. On these recording days, rats had either free access to food and water rewards (sated), or only a limited access to both, making both food and water either equally preferred, or equally non-preferred. When examining the factor 'Goal' on these days, the proportion of cells encoding either the food or water port, prior to entry, remain virtually the same (Figure 8D). There were still significant proportions of neurons differentiating between the two ports on these days. Interestingly, when examining the Identity index, the only result that remains the same is that the Shell still shows a bias for the Water Identity (Figure 8D inset). The DMS, DLS, OFC and Core all have equal numbers of cells firing for both food or water reward, showing no specific Identity bias. This means

that there is no longer a preference for the food reward in the DMS and DLS.

This throws a wrench in the theory that the encoding seen for Goal Port encoding was due ONLY to the fact that it was reward identity, with no influence of value. There now appears to be some change when the rat is hungry or thirsty, compared to when he is only mildly food/water restricted or sated. The possible explanation for this is that the rat had a tendency to view the food reward as more valuable, overall, on days where he was food or water restricted. When he was no longer specifically deprived of food or water, that signal went away, and there was no longer any bias in the regions that had previously signaled a bias for the food port.

This explanation could be due to an effect of general motivational state on reward responding. If the rat was in a motivationally charged state (food or water restricted), than the brain may signal that food = better, even on water restricted days. On sated days, both water and food had an equally low value, and so equal numbers of cells differentiated between the food and water ports, but with no bias towards identity or value. In a motivationally charged state, the food port gains greater value for the brain regions influenced by (possibly) a dopamine signal, and therefore the DMS and DLS increase response for the food port, as identity, but this can only happen when the rat is more motivated, and not on sated and restricted days.

Model-free vs. Model-based Encoding Schemes

A very popular concept in striatal research right now is to examine neural encoding in specific regions through the lens of reinforcement literature. Within the current understanding, different areas of the brain operate in separate ways when

representing what the rat intends to do, as far as decision-making and action selection go. In a 'model-based' representation of the world, the specific stimulus of interest is represented in a rats mind, and he acts accordingly. In a 'model-free' environment, the rat does not have a specific goal in mind, and instead executes an action that is known to be good due to prior experience, but without representing the specific outcome. In the model-based framework, the rat has to think and plan ahead, using a search like option, to make a decision, which takes time and energy. On the other hand, time is saved by using a model-free framework by 'caching' the values of specific actions, and carrying out the decision without searching for the representation of the upcoming goal.

Based on previous lesion work, the DMS is thought to operate in a model-based environment, encoding action-outcome contingencies and representing the goal in mind. The DLS is thought to operate in a model-free environment, in a more habitual manner, without prior goal-orienting knowledge (Daw et al., 2005). Single unit recording studies seeking to specifically test these hypothesis using tasks that utilize both a model free and model based encoding of a task have not found the precise neural correlates in the rodent striatum (Stalnaker et al., 2010; Thorn et al., 2010).

The results from this study show that there is most likely model-based encoding in all areas of the striatum, OFC and nucleus accumbens. As the rat approaches the reward port, specific goals can be distinguished in the neural firing rates of all 5 brain regions, beginning up to 2 seconds prior to actual port entry. If the idea of model-free vs. model-based decision-making is based on whether or not a specific goal-identity is in mind, then a direct comparison of two distinct goals directly tests this. In this task, the rewards are food and water, and through the rat's behavioral choices, it is obvious that

one reward is more preferred, depending on the hunger or thirst of the rat. However, since the rat must go to a specific port in order to actually collect the reward, this study directly tests if these goal representations are 'in mind'. If the rat did not have a specific representation of his intentions, then there would be no differentiation in firing rate as the rat approached the reward port. This would be consistent with a model-free form of encoding- there could possibly be an increase in overall activity as the rat approaches the reward, but an actual differentiation between the two rewards would be unlikely, since that would indicate some type of distinct representation of goal identity. Another possible encoding of a model-free representation would be firing for the specific Cue, or stimulus, that played at the beginning of the trial. This signaled the correct reward port, and could have been represented in a stimulus-response manner, without actually having a goal associated with it. The Cue simply represents a response, such as 'go to port on left'.

Within this study, if the DLS had a model-free representation, it would not encode a significant proportion of cells differentiating between the two specific reward ports. This is NOT what was found. Instead, each of the 5 brain regions examined had distinct populations of neurons that fired differently, based on if the rat was approaching the food port or the water port. Having a specific representation of any type of goal, prior to actually realizing it, is consistent with a model-based theory.

Current theories and previous lesion studies have found distinct and separable roles for various BG loops in processing action-selection, while single-unit studies have found conflicting results. The results from this behavioral task, which directly tested 2 distinct types of rewards, with different preferences based on internal state (a more natural representation of value), provides insight that rather than parallel loops carrying

distinct information through the basal ganglia, a more integrated system is present. This integration includes a representation of the different components of a decision-making task, including representing a specific goal in the DMS and DLS, as well as nucleus accumbens and OFC.

Reward Port Encoding Conclusions:

The overall conclusion from this analysis on the factor 'Goal' is somewhat surprising, given past studies. First of all, the fact that there was any differentiation at all between two natural reward identities was unexpected. Since most other studies simply examine the anticipation of one rewarding event, it seemed like the signal from the striatum was a generally motivational, or generally reward-predictive signal. However, from comparing 2 separate rewards, both of which are valuable, it is clear that the specific identity of that reward is actually differentiated in the brain. If this were related to the fact that one of the rewards was more valuable than the other, then the findings here would be directly in line most other results. However, the results from this study show that specific differentiation does NOT depend solely on internal state, and that the identity of the upcoming reward is encoded, with dorsal striatum showing a bias for the food identity and the nucleus accumbens showing a bias for the water identity. When the rat was no longer in a specifically food or water restricted state, all regions still differentiated between the two Goals, but identity was only encoded in the Shell for the water reward. This could mean that food port identity is only encoded in the DMS, DLS and OFC when the rat is in a state of higher general motivation, which then signals a stronger bias for the food Goal.

These results make it more difficult to fully understand how previous reward-related firing in studies that only study one reward truly show 'reward anticipation', or predict a probability of reward. If different reward types are not being compared, other studies can not rule out reward identity as the corresponding neural signal.

In contrast to the Goal only cells, there were no regions that showed significance for the actual cued reward before the rat got to the reward port (Figure 9A). This result is somewhat surprising, since we expected to see encoding of the tone that played in at least one region, most likely the Shell, Core, or DLS, in order for the animal to keep the Cue in mind. The Cued reward was the tone that played, which corresponds to the correct port the rat had to enter. If a brain region encoded the Cued reward, rather than the Goal, then it would be signaling the 'rule' that the rat needed to follow, and could be interpreted as a signal that overcomes the Pavlovian instinct to always pick the more preferred reward port. By utilizing the multiple regression analysis that had factors of Cue vs. Goal, we can be sure we know what, specifically, the cells were firing for. It turns out the prominent signal before the rat entered the reward port was for Goal, with no regions reaching significance for Cue. After reward port in, there were cells that reached significance for the 'Cue' factor, but they also reached significance for the 'Goal', as well as an interaction between the two- these will be discussed later.

Research from the Graybiel lab (Thorn et al., 2010; Barnes et al., 2005; Barnes et al., 2011) are some of the only other rodent studies that specifically instructed the rats about what the upcoming choice should be, in a task where the rats still had to make a choice, and were wrong in some trials. In their study, unlike the current study, they did not see any Cue encoding at the time of tone presentation. They also did not see any

encoding of the tone/cue that played when the rat actually made the decision. The results from the current study add further proof that neither the striatum, nor the OFC, encode information about the Cued stimulus prior to the actual decision, despite the fact that we expected to see some encoding of the correct answer prior to reward port entry.

Outcome encoding was another significant finding that was shared across each brain region. Finding a specific signal related to receiving a reward is not novel. Most research in the striatum in rodents has found an outcome signal in both dorsal and ventral striatum (Oyama et al., 2010; Kim et al., 2009). Comparing the DMS and NAC, Kim 2009 found a signal related to the current outcome just after a choice was revealed. The NAC is often shown as the region that encodes whether or not a reward is or is not received which is often used as evidence to support a role for the NAC as a 'critic' (Kim et al., 2009; Taha and Fields, 2005; Van Der Meer and Redish, 2009a).

Rodent OFC studies have also found outcome related firing (Sul et al., 2010; Feierstein et al., 2006; Furuyashiki et al., 2008; van Duuren et al., 2009). Some studies have even found firing for the upcoming outcome before it's actually revealed in the rodent OFC (Kepecs et al., 2008). Sul et a.l found that firing in the OFC, after a choice was made, had a combination of factors related to the chosen value, including expected outcome (Sul et al., 2010).

Interestingly, the DLS is the only region that fires for the correct reward, while all of the other regions fired much more for the incorrect reward. However, the DLS switches this preference 2.2 seconds after reward in, and begins firing more for incorrect reward. At this time, all 5 brain regions are still encoding a strong Outcome signal. One possibility is that the Core, Shell, DMS and OFC are encoding some variant of a reward

prediction error- that the rat truly believed he was supposed to receive a reward, despite having gone to the wrong port, and is firing for the missing, but expected reward. The DLS may be the region that knew the answer all along, and is integrating the information of what was cued, what was chosen, and what was received in order to 'stamp' that into memory, and signal a positive outcome. The regions firing for the incorrect reward may be doing the opposite, firing when a reward was not received to remember that result into the next trial, and not repeat the same mistake twice, since the trials are repeated when wrong.

Another possibility is that the DLS is only responding to the motor movements required to consume the reward at the 700ms time point, then switches over to more cells encoding the incorrect reward in a similar manner as the other 4 regions. This would only be if the DLS motor movements were very general- that is, the same for food and water. This is because the current outcome factor only takes into account whether a reward was presented or not- it does not matter if it was a food or water reward. The motor movements of chewing vs. licking are quite different, and it would make sense that they would actually use different muscle groups encoded in separate areas of the DLS.

The findings in this study show that there is a separation in signals between specific aspect of reward identity and encoding for receiving the reward. We can directly compare the signal that relates to specifically receiving a reward, and the signal that relates to the goal. By having a behavioral task with longer time intervals between events, and utilizing two different reward identities, we can see that prior to entry into the reward port, the neural signal encodes the identity of the upcoming reward ('Goal'). After the rat either does or does not receive a reward, a separate signal arises that encodes reward

receipt ('Outcome'). The different signals we see in this analysis most likely overlap with previous results seen in these brain regions in signals that are interpreted as reward anticipation, or an outcome signal. Obviously, it is impossible to fully probe exactly what the rat was thinking before he entered the reward port- whether or not he knew he would be receiving a reward- but we do know that receiving or not receiving a reward was a different signal than the integration of getting a specific reward, which will be discussed in the next section.

The final main result that came from the 3-factor analysis was for cells that had 2 main effects + an interaction- for Cue, Goal and their interaction (outcome). Like the results seen for Outcome, the cells that showed significance for reward integration only did so after the outcome of the trial was revealed. The DLS had the highest proportion of cells showing this integration effect, and again, like the Outcome results, the DLS as a region was more biased for the correct trials, while the other 4 regions showed a significant bias for the incorrect trials.

This combination of terms is the only other instance, besides action-goal, when a brain region shows value-related firing, rather than identity. In the DMS, and the DMS alone, the cells that are significant for reward integration show a bias for the non-preferred reward port (they did not show significance for the identity). The fact that the DMS is the only region to show a value-related bias, and only does so in cells that integrate more than one factor (action earlier in the task, and goal after reward port in), including the interaction effect, is an interesting result. It also just so happens that the DMS cells are doing so in the same direction in both analysis- for the non-preferred port. There is a group of DMS cells that fire more when the rat is selecting the non-preferred

port, and a group of DMS cells that encode information about the non-preferred trial upon trial completion. However, these same cells are not showing a bias for correct vs. incorrect- they are firing differently for it, but not in a biased manner. The DMS cells that are integrating information after reward port in about the non-preferred reward port entry are doing so significantly on both correct and incorrect trials.

This integration of Cue, Goal and Outcome is the only other time, besides at the time of the tone, when the factor 'Cue' reaches significance. Just like at the time of the tone, the DLS stands out in this combination of factors, having a much greater proportion of cells with the 3 effects than all of the other brain regions. Interestingly, the DLS shows a bias for firing the most for correct trials right after reward port entry.

The integration of Cue-Goal-Outcome may actually be a reward consumption signal. The rat most likely uses different muscle movements for consuming food vs. water, so depending on which reward the rat received would dictate the motor aspect of consuming the reward (chewing vs. licking). The reason there is a main effect of Cue AND Goal is due to specific differentiation between actually receiving one of the two rewards. For example, a licking cell would show a response for water cued AND water chosen to actually receive the water reward (outcome). On the other hand, if there was only a main effect for Cue + Outcome, this would mean that the rat differentiated rewarded vs. no reward, but only did so on water-cued trials, regardless of where the rat went, and therefore would not differentiate a licking movement. A main effect of Goal + Outcome would mean that the rat differentiated reward vs. no reward, but only for when the food port was chosen. However, having a main effect of Cue + Goal, + Outcome means that the rat had to choose the cued reward and then differentiated reward vs.

unrewarded, which could lead to a specific motor movement signal.

Reward History Encoding Across Regions

The combination of factors seen in Figure 13A and 14A shows that both the Core and Shell prominently encode reward history. As the outcome of the current trial is revealed, both Core and Shell fire more for the incorrect reward (Figure 13B). In the 3 seconds before the tone plays, only the Core fires significantly more for the previously incorrect trial (Figure 14B). Although all 5 brain regions are encoding Previous Outcome, there is only significant bias in the Core. As the rat continues the current trial, signal related to the previous outcome begin to fade in all regions except for the Shell. The Shell only briefly shows a bias for the previously incorrect reward, but it is only for 500ms.

This type of firing rate pattern has been seen before, in Kim et al. papers- where reward history influenced firing rate both in the intertrial interval and into the next trial. Previous research finding reward history has been seen in the rodent OFC (Sul et al., 2010; Young and Shapiro, 2011; Takahashi et al., 2011; Roesch et al., 2006), the rodent NAC (Kim et al., 2007; Goldstein et al., 2012), and primate striatum (Hori et al., 2009; Yamada et al., 2007; Yamada et al., 2011).

Finding information related to previous outcome is not ubiquitous among regions. Thorn (2010), directly compared DMS and DLS and did not find encoding of the previous outcome in either region. Contrary to that, other rodent studies have found previous outcome encoding in the DMS (Kim et al., 2009; Kim et al., 2013; Kimchi and Laubach, 2009a), though previous outcome finding in the rodent DLS is relatively sparse (Kim et al., 2013). Interestingly, most previous research that found a signal related to

previous outcome has not specifically analyzed (or presented) what that information was. It may be due to the varying types of tasks used. This gives my analysis a unique contribution, by differentiating what the signal conveys from the previous outcome.

In the current study, for the duration of the next trial, the two regions of the ventral striatum encoded information about the previous trial the longest. Both Core and Shell had a very strong encoding about whether or not the previous trial was rewarded up until the rat received a reward in the current trial. The Shell held onto the previous outcome information the longest. Additionally, at the time of the tone, the Core showed a significant bias for the previously incorrect outcome- meaning the cells were firing more on trials where the rat had previously been wrong, therefore making the current trial a 'repeat' trial. The Shell, OFC, DMS and DLS did not show a bias in the cells that encoded previous outcome, though each region did significantly encode this factor.

Additional regression analyses were run on other aspects previous trial encoding, based on the Cue or the Goal, but no brain regions reached significance for any single factor or combination of factors. This is interesting because previous outcome is even more informative in this behavioral task than most- it actually could determine the trial type in the next trial. If the rat picked the incorrect reward port, he had to repeat that trial, and that gave the rat more information going into the next trial than he would have if he had picked a correct trial. On correct trials, the probability of preferred vs. non-preferred reward goes back to 50-50. The fact that the NAc regions encode this previous outcome factor for so long into the next trial seems to show that the NAc may actually be performing as an informative critic, by utilizing information from the previous trial.

However, since there is no interaction between other terms in the current trial, it isn't clear exactly how the NAC is influencing the current trial.

Kim (2007) found strong encoding of information related to the previous outcome in the ventral striatum. Their behavioral task was similar to this task and did not require an actual choice. Rats were always rewarded with a constant amount of water for visiting the lit side of a figure-8 shaped maze. This was a visual discrimination task- they were only rewarded when they visited the lit side of the maze. Interestingly, the rats only performed correctly in >70% of trials. They did find activity in NAC for the animals' choice during reward approach, reward consumption, and return (after the reward). However, prior to the choice, or response selection stage, less than 3% of neurons were modulated by the upcoming choice in the current trial, while at the same time, 17% of neurons were modulated by previous choice. Although the current study shows significantly more cells firing for the upcoming choice than they do, that difference is most likely due the difference in tasks. There is a similar pattern in NAC activity about the previous outcome all the way up to the Goal port choice in both tasks.

Another study that found encoding for the previous outcome was in the DMS and DLS in Kim (2013). This finding has not been seen quite as prominently in the DMS/DLS, so the findings in the current study, coupled with the findings in Kim (2013) show a more interesting role for the dorsal striatum in decision-making. Previous research and theories have tried to contend that the NAC and DS have separate roles as an actor or critic, based on the type of information and timing of those information signals. However, the strong overlap in all of these regions, especially for the previous outcome signal, calls into question the role of the NAC as the only area that can act as a

critic. In Kim (2013), the previous goal choice modulated neural firing rates for a longer time in the DMS, but was still there in the DLS. Interestingly, they found that there were higher firing rates for unrewarded trials, but this was only seen as a network pattern, rather than a signal that was carried by a few neurons. In the current study, there was not a strong bias for previously unrewarded firing in the DMS, but instead, this signal was found in the Core. It may be a difference in the nature of the task- they used a probabilistic free choice task so the rats did not expect to get a reward on every trial, while the signal the rats had in the current study was a more direct error signal.

In this way, the information processing done by the DMS and Core may be able to be distinguished. Although neural representation is strongly overlapping in all striatal subregions, the exact content of that representation is more differentially modulated by the precise nature of the task. In a more unknown environment, one without specific Pavlovian associations, like the free choice probability task used by Kim (2013), the DMS encodes the unrewarded outcome. In a cued-task, with a Pavlovian association that also involves an instrumental response, the Core may reliably encode the unrewarded outcome.

In the Core, there is a strong bias for previously incorrect trials, however the multiple regression for previous outcome + current outcome did not show any significant interaction effects among cells, so it appears as though the cells that encode information about the previous outcome are separate from those that may encode information about the upcoming outcome, or upcoming chosen reward. What are these cells encoding, if they aren't actually informing the behavior of other cells? It is possible that increased firing after incorrect trials in the Core leads to increased vigilance and motivation into the

next trial, so the rat can get the next trial correct. The fact that the other regions also modulate their firing rate for previously incorrect trials shows that this information is informative, but it is puzzling that no other regions show a bias in firing between the previously correct vs. incorrect. It seems like this information would be particularly informative after incorrect trials, because those trials are repeated, so in order to remember that, or be more vigilant, firing should be higher on previously incorrect trials. Instead, all other brain regions have even encoding of both trial types. To determine the exact nature of this encoding is beyond the scope of this project, but would be extremely informative, since Previous Outcome is another factor where clear regional differences were found.

This behavioral task is unique in the literature on decision-making because the behavior and subsequent analysis utilized both correct and incorrect trials. The fact that the rat's behavior was steady, and their errors were not random (ie, they were nearly always in the direction of the preferred port) indicates the rat understood the stimulus-response pairing, but was often compelled by his internal state to seek out what he wanted, and incur an error/time out. Other behavioral tasks that have tested reward and outcome-related encoding have rarely used a task where reward was instructed, but animals chose to perform in an incorrect manner. Other paradigms use probability of reward at a specific port and a free choice paradigm to look for outcome or goal in different brain regions. This study directly compared where the rat was instructed to go, with where he actually went. This direct comparison leads to critical insights in instruction vs. intention, and provides interesting ideas about what happens when a conflict arises. The results show that all 5 brain regions care most about where the rat

actually goes, regardless of the instruction. None of the 5 brain regions reliably encode the instructed port during performance of movement. Additionally, despite each port having specific value depending on internal state, no brain region was biased to fire more for the more highly valued reward.

The brain regions studied all clearly identified a difference between the choices the rat was about to make. If this signal was to fit into a distinct model-free vs. model-based encoding framework, we would expect to see regions that never reached significance between the different reward port identities. If that had been the case, then the response the rat was engaged in could be seen as potentially habitual, and lacked specific goal-related encoding. Another piece of evidence for a model-free encoding result would be if the rat integrated information about the stimulus, or cue, prior to reward entry. In this case it would be indicative of a representation based on something other than the specific reward identity and it could be argued to be a stimulus-response, or habitual action. This is not what was seen. All 5 brain regions had specific populations of neurons that fired differentially between food and water ports, showing clear model-based encoding prior to reward port entry. The first conclusion in this analysis shows that all 5 brain regions encode reward choice by specifically representing the Goal.

The second conclusion is based on a test of value vs. identity. It has been proposed and shown that firing rates seen before receiving a reward are an anticipatory signal based on an estimation of the value of that reward. If this were the case in the current study, then internal state, which influences behavior, would subsequently influence the firing rates of individual neurons and lead to a higher firing rate on preferred reward port entry. Again, this is not what was found. If the regions had no bias

about value, and instead were firing equally for preferred vs. non-preferred reward, then we would also expect to see no bias when testing for identity, either. The analysis for reward identity DID show a significant bias in brain regions- food in the dorsal striatum and OFC and water in the NAc. This shows that the population of neurons that differentiate between the 2 reward ports do so based on the specific goal identity rather than value. Although this has been demonstrated in the past, it is a finding that is not often probed, or considered, when examining neural firing rates that are presumed to underlie 'value', mainly because studies typically examine only 1 type of reward.

A third important finding from this analysis showed the differences in timing and regional biases related to outcome. This signal- whether or not the trial was rewarded- was significant in all 5 regions. The NAC is often thought of as the 'Critic', showing the greatest signal about whether or not a trial was rewarded, but other studies have found similar results in the Dorsal Striatum, calling into question the specific contribution the NAC has in decision-making. By directly comparing these two regions, this study shows that they both have distinct cells that specifically fire based on whether or not a reward was delivered. These cells independently encoded this factor from other potential neural correlates such as action, the actual goal port the rat went to, or the cue that played, and instead was solely based on receiving the reward or not.

Another important results from the current outcome factor showed a significant difference in the processing of outcome. The DLS stood out as the only region that increased firing for correct rewards. The fact that incorrect rewards had higher firing rates in DMS, Core, Shell and OFC could be seen as a negative prediction error, while the

increased firing in DLS could be seen as a positive prediction error. However, the behavioral task was not designed to specifically test this, so this is only speculation.

Before the rat entered the reward port, there were no brain regions that demonstrated knowledge of what the outcome of the trial would be. No region fired for the correct cue, or the current outcome, or any integration between cue/goal/outcome. After the rat received the reward, there was finally an integration of all factors related to reward. The timing of this onset, though, could reflect consumption of the specific reward- that there are specific cells in all 5 regions that are actively engaged in only eating vs. drinking, based on the different motor movements involved.

The final result from studying activity at the reward port is the relative contributions of the regions. Each region tended to show a relatively different contribution to each of the factors at reward port in. The DMS, which had higher firing during action-goal, showed relatively little firing rate differences for the upcoming goal, comparatively. The Core and OFC had the highest contribution to the actual Goal. The Core had the highest contribution for the Outcome, while the DLS showed the highest proportion of cells integrating all of the factors at the reward port.

The conclusion is that all 5 brain regions have an overlapping representation of all aspects of specific reward coding, with a model-based, identity-based encoding scheme. None of these findings are actually novel, but shown together, in the same task and in the same rats, underscores the overwhelming overlap in signals in all of these regions, and argues against very specific regional encoding that is left to distinct subregions. There were, however, relative differences in proportions from each region, which could explain some of the behavioral differences seen in lesion studies.

Another potential explanation that could explain the gap between lesion and single unit studies is that the initial contribution of each factor may originate in a specific locus, which is then sent to the other brain regions to become integrated with the signals that originate in that region. Since the basal ganglia is shown to have loops that integrate information between specific subregions, it could be that these results are the reverberation of signals accumulating throughout the regions, as they overlap in the informational processing network. Despite having precise temporal resolution in single unit studies, it may be that the network timing is even faster than can be detected with the current data analysis, so the net result is a recording of the integration of all signals in the BG network after they were initiated and shared.

Putting it all together: A shared framework with specializations

The results from this research project give answers to some questions in the field of basal ganglia research, while also opening new and exciting avenues of research about how an animal makes a decision. The first contribution this study makes to the field is the introduction of a unique behavioral task that has the ability to separate and identify specific aspects of decision-making- the Action, Goal, Cue, Outcome and the integration of these factors. Further research can be done probing various aspects of this internally-driven behavioral task, both through behavioral studies, lesion studies and more electrophysiology studies. The reaction time differences seen in some of the individual rats open questions about how motivation may influence both reaction and movement times. Questions about movement times related to errors can also be studied behaviorally as well as with single unit recordings. Even studies testing learning can probe how a rat's behavior and reaction times track with his choices as he learns the meaning of the different tones throughout training. Lesion studies should also be undertaken to fully understand each regions unique contribution to the different aspects of the task- in both performance and acquisition of the behavior.

Questions that were answered in this study relate to the specific contribution that individual neurons have within a region that participates in decision-making and reward-related behavior. One question that is seemingly answered with lesion studies, but is not

as well understood in electrophysiology, is the contribution of the DMS and DLS to habit vs. goal-directed behavior. On one hand, previous lesion work in rodents has demonstrated strong regional differences between the DMS and DLS in the acquisition and performance of instrumental behavior. On the other hand, single unit recordings attempting to discover the exact neural underpinnings of those differences have not replicated the clear distinctions seen in the lesion work. By directly comparing these two subregions in a novel behavioral task that used a more indirect measure of value- internal state- we can address the similarities and differences.

Based on lesion studies, the DLS should be participating in this task in a habit like manner, since the rats were highly overtrained in the task, and therefore should show very little influence of internal state on the neural firing rates. It would also make sense to see more stimulus-response coding in the DLS, since it is proposed that this is how the DLS processes more habitual behavior. Additionally, connections between the DLS and motor cortices would indicate a greater involvement in the motor aspects of the task, compared to DMS and the nucleus accumbens. Since the DLS is thought to encode information in a model-free manner, there should be little influence of internal state, and potentially little to no encoding of the separate rewards, since that would mean that the information was being represented differently, based on the identity of the outcome.

In the DMS, lesion studies predict that the involvement of this region would be in encoding the action-outcome association. Models of the DMS using reinforcement literature theorize that the involvement of the DMS is in representing information in a model-based way, so the neural coding of the region would represent the identity or value

of the chosen reward. The hypothesis would be that the DMS was representing value while working for reward.

The results of this study both agree and disagree with these hypotheses. The first interesting result is that there was a surprising amount of overlap in the encoding profile seen between these two regions. That is, both areas of the DMS and DLS encode a great deal of information about the Direction of movement, the Goal the rat was working for, and the Outcome of whether or not the rat was correct. Of particular surprise was the fact that there was almost equal encoding of Action in both regions.

The main differences in encoding schemes between these two subregions offer some support to current theories on how information is processed in these neural networks. The Cue is the first instance where sub-regional differences show up. As predicted, the DLS encodes information about the Cue, or stimulus, while the DMS does not. This indicates that the DLS may be responding to the stimulus in a habitual, or stimulus-response like manner. Additionally, the DLS is not encoding this information with respect to value, since it is responding equally regardless of if the tone presented indicates a more or less preferred reward. This difference between DMS and DLS in encoding for the Cue is the first instance that shows different contributions to the overall encoding schemes in a decision-making task.

The second instance where sub-regional differences manifest is when the rat is performing the action. Both regions show nearly equal proportions of cells that differentiate between the contralateral and ipsilateral movements. The difference is in the nature of how that information is encoded. The DMS fires more on the contralateral trials, while the DLS fires equally for either direction. Despite the more direct connection

with the motor cortices projecting to the DLS, it appears that both regions fire for motor movement. The differences between the regions may be due to the specific connections and composition of each region. The DMS may have a more intricate connection with the precise motor movements of moving to the left or right, therefore activating more direct pathway neurons and showing a bias for the contralateral direction. On the other hand, the DLS may have a more broad, generalized encoding scheme and therefore does not activate differences between the direct and indirect pathway.

The third instance of sub-regional difference is seen as the rat moves to the final nose poke. In this instance, the DMS is encoding a combination of the contra/ipsi directional movement while also integrating information about the final goal port. The DLS is not integrating this information. This encoding is exactly what we hypothesized for the DMS, an action-outcome like association. In this case, the 'outcome' is actually the goal port the rat will then enter. Additionally, the DMS is encoding this information based on value, and is the first instance when value based on internal state actually manifests in this task. Finally, after the rat reaches the goal port and either did or did not receive a reward, there was another difference in encoding between the DMS and DLS. Both regions showed a difference in firing rate between correct and incorrect trials, but the DLS fired more for the correct reward, while the DMS fired more for the incorrect reward.

These sub-regional differences in a single unit study show that some ideas from lesion studies about the relative contributions to reward-based behavior in the DMS and DLS have distinct neural underpinnings. The DLS may respond specifically to a Cue in a more stimulus-response like manner, and not encode this information based on value. The

DMS does integrate action-outcome information, and does encode this information based on value, albeit by firing more for the non-preferred reward.

Prior studies in the Accumbens have shown that the Core and Shell may integrate motor and reward information. Other theories show the NAC responding to cues that predict reward. More theories hypothesize a role for the NAc as a critic that informs an animal about his decision, rather than actually making that decision. This study directly compared both Core and Shell in a task that separated Action from Goal, and could specifically test the contribution of the NAc in direct comparison to other areas of the striatum.

In this study, there was a great deal of overlap between the Core and Shell. First of all, both regions did show specific responses to the movement (contra vs. ipsi), but did not reach as high of proportions as the Dorsal Striatum and OFC. It was surprising to find, however, that at the time of movement, neither Core nor Shell integrated information about this movement with any other aspect of the task. Additionally, both regions showed a similar response to the Goal port as the rat entered to get a reward- both Core and Shell fired more for the water port.

The one major difference between the two regions was the opposite of our hypothesis. The Shell responded more to the Cue than the Core. The only other major difference was the fact that the Core fired more for the previously incorrect trials, while the Shell did not show this same bias. Both regions maintained firing for the previous trial well into the next trial.

Another difference between the Core and the Shell shows up when examining how cells that encode information overlap between different events (Figures 15-24). The

Core has a great deal of overlap between cells that encode information between events (Figures 17 and 22). Specifically, there are very few cells that only encode action during movement. Most of the cells that respond to the two different directions also respond to either the outcome or the Goal. In contrast, the Shell, as well as both areas of Dorsal Striatum, have more cells that encode only the Action, with fewer cells overlapping between events (Figure 18). The Core is encoding information about the action in the same cells that go on to select the goal, or encode information about whether a reward was collected or not. The Shell and Dorsal Striatum seem to separate these aspects of information into separate cells at separate time points.

In all regions, including the OFC, the overwhelming majority of cells encode Goal, with far more cells that only encode this information, without overlap between Action, Outcome or Integration of that information. This is important information for future studies that are only testing one type of reward. It is apparent that a great deal of individual striatal cells encode specific information about identity, not value, or simply a general reward signal, but this signal is usually not tested for identity.

The final contribution of this study is answering questions that relate to value vs. identity. Most single unit studies currently use only one form of reward when testing decision-making. Neural firing rates are then analyzed and assumptions are made about value, without always considering identity. Since this study used 2 different types of rewards that changed value based on internal state, it was a direct comparison of value vs. identity. The results show that all 5 brain regions most often encode identity.

Breaking down the regions, each subregion has a unique profile that contributes to the overall decision-making network:

Core:

The Core's highest contribution is encoding the Current Outcome after reward port in (Figure 13A and 17). The Core fires most for the incorrect reward, and continues to encode this information until the tone plays in the next trial. Within the Core, though, the highest proportion is Goal (Figure 17A).

Shell:

The Shell's contribution is to have the highest encoding for the Water tone, relative to the other brain regions (Figure 9A and 9C). Within the Shell, the Goal reaches the highest encoding, and it does so for Water, on all testing days (Figures 8C and 12A and 12C). The Shell also encodes Reward history for the longest period of time, doing so in a select group of cells that appear to encode this information across multiple events (Figure 23).

DMS:

The DMS has the highest action-goal cells, possibly as a role in action selection on non-preferred trials (Figure 11). The DMS also integrates information about non-preferred trials after the outcome is known to the rat (Figure 13C and 13D).

DLS:

The DLS is the odd-ball out. When all other regions fire more on contralateral trials, the DLS does not have a bias (Figure 7C). All other regions fire more on the incorrect trials, while the DLS fires more on correct trials (Figure 13B). As far as proportions go, the DLS as a region fires most when integrating Cue + Goal + Outcome, and which is most likely specific motor movements for eating vs. drinking, by firing most on the correct trials, regardless of preference (Figure 13D and Figure 15).

OFC:

The OFC has the highest proportion of cells for both Action and Goal (Figure 7A, Figure 12A and Figure 19). It seems to contribute to everything, with only the Cue not showing significance. It also shows the most overlap between different events and different encoding factors, showing that individual cells change their encoding scheme based on the current behavioral event.

Figure 25: Overall Encoding

Each region of the Basal Ganglia and OFC has a specialized contribution to translating Internal State into Action Selection

In an overall scheme of encoding, the idea that the basal ganglia have integrative loops that possibly share information across regions is the best way to explain the current results. Lesion studies clearly show that when a region is inactivated, specific aspects of behavior are lost. In the current study, there were only a few instances of a region being the sole contributor to a specific aspect of behavior. Instead, the results show that all of the regions do actually encode information about nearly every aspect of the task, but the relative contribution of each region changes, depending on a specific aspect of behavior. The process of translating internal state to reward-guided decision-making is achieved through a network of shared information throughout integrated loops in the cortical-basal ganglia network. Each region is responsible for a main component of the decision-making process, and then shares that information throughout the network. This information is highly distributed between individual cells within each region, since there are only a few cells that encode information for a lengthy period of time. Information processing happens through a distributed network of encoding- between cells, and between subregions- in order to decide to head to the cupboard, select the blue and white Nabisco package, and eat the Oreos. Got milk?

Bibliography

Abe, H., and Lee, D. (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. Neuron *70*, 731-741.

Aberman, J.E., and Salamone, J.D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. Neuroscience *92*, 545-552.

Albin, R.L., Young, A.B., and Penney, J.B. (1989). The functional anatomy of basal ganglia disorders. Trends Neurosci *12*, 366-375.

Alexander, G.E., Crutcher, M.D., and DeLong, M.R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. Progress in brain research *85*, 119-146.

Alexander, G.E., and DeLong, M.R. (1985). Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. Journal of neurophysiology *53*, 1417-1430.

Alexander, G.E., DeLong, M.R., and Strick, P.L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. In Annual review of neuroscience, pp. 357-381.

Ambroggi, F., Ghazizadeh, A., Nicola, S.M., and Fields, H.L. (2011). Roles of nucleus accumbens core and shell in incentive-cue responding and behavioral inhibition. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 6820-6830.

Ambroggi, F., Ishikawa, A., Fields, H.L., and Nicola, S.M. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. Neuron *59*, 648-661.

Apicella, P., Scarnati, E., Ljungberg, T., and Schultz, W. (1992). Neuronal activity in monkey striatum related to the expectation of predictable environmental events. In Journal of neurophysiology, pp. 945-960.

Aragona, B.J., Day, J.J., Roitman, M.F., Cleaveland, N.A., Wightman, R.M., and Carelli, R.M. (2009). Regional specificity in the real-time development of phasic dopamine transmission patterns during acquisition of a cue-cocaine association in rats. The European journal of neuroscience *30*, 1889-1899.

Atallah, H.E., Lopez-Paniagua, D., Rudy, J., and O'reilly, R. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. In Nature neuroscience, pp. 126-131.

Balleine, B., and Killcross, S. (1994). Effects of ibotenic acid lesions of the nucleus accumbens on instrumental action. Behav Brain Res *65*, 181-193.

Balleine, B.W., Delgado, M.R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. The Journal of neuroscience : the official journal of the Society for Neuroscience *27*, 8161-8165.

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

Balleine, B.W., Liljeholm, M., and Ostlund, S.B. (2009). The integrative function of the basal ganglia in instrumental conditioning. Behav Brain Res *199*, 43-52.

Balleine, B.W., and O'Doherty, J.P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *35*, 48-69.

Barnes, T., Kubota, Y., Hu, D., Jin, D., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. In Nature, pp. 1158-1161.

Barnes, T., Mao, J.B., Hu, D., Kubota, Y., Dreyer, A.A., Stamoulis, C., Brown, E.N., and Graybiel, A.M. (2011). Advance cueing produces enhanced action-boundary patterns of spike activity in the sensorimotor striatum. In Journal of neurophysiology, pp. 1861-1878.

Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. Nature neuroscience *7*, 404-410.

Bassareo, V., and Di Chiara, G. (1999). Differential responsiveness of dopamine transmission to food-stimuli in nucleus accumbens shell/core compartments. Neuroscience *89*, 637-641.

Basso, A.M., and Kelley, A.E. (1999). Feeding induced by GABA(A) receptor stimulation within the nucleus accumbens shell: regional mapping and characterization of macronutrient and taste preference. Behav Neurosci *113*, 324-336.

Bauter, M.R., Brockel, B.J., Pankevich, D.E., Virgolini, M.B., and Cory-Slechta, D.A. (2003). Glutamate and dopamine in nucleus accumbens core and shell: sequence learning versus performance. Neurotoxicology *24*, 227-243.

Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron *47*, 129-141.

Bennett, B.D., and Bolam, J.P. (1994). Synaptic input and output of parvalbumin-immunoreactive neurons in the neostriatum of the rat. Neuroscience *62*, 707-719.

Berke, J.D., Breck, J.T., and Eichenbaum, H. (2009). Striatal versus hippocampal representations during win-stay maze performance. Journal of neurophysiology *101*, 1575-1587.

Berke, J.D., Okatan, M., Skurski, J., and Eichenbaum, H.B. (2004). Oscillatory entrainment of striatal neurons in freely moving rats. Neuron *43*, 883-896.

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

Bertran-Gonzalez, J., Bosch, C., Maroteaux, M., Matamales, M., Herve, D., Valjent, E., and Girault, J.A. (2008). Opposing patterns of signaling activation in dopamine D1 and D2 receptor-expressing striatal neurons in response to cocaine and haloperidol. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 5671-5685.

Bornstein, A. M. and N. D. Daw (2011). Multiplicity of control in the basal ganglia: computational roles of striatal subregions. Curr Opin Neurobiol **21**(3): 374-380.

Bouret, S., and Richmond, B.J. (2010). Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 8591-8601.

Bowman, E.M., Aigner, T.G., and Richmond, B.J. (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. In Journal of neurophysiology, pp. 1061-1073.

Brasted, P.J., Humby, T., Dunnett, S.B., and Robbins, T.W. (1997). Unilateral lesions of the dorsal striatum in rats disrupt responding in egocentric space. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 8919-8926.

Bromberg-Martin, E., Matsumoto, M., and Hikosaka, O. (2010). Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. In Neuron, pp. 144-155.

Brown, V.J., and Robbins, T.W. (1989). Elementary processes of response selection mediated by distinct regions of the striatum. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 3760-3765.

Bryden, D.W., Burton, A.C., Kashtelyan, V., Barnett, B.R., and Roesch, M.R. (2012). Response inhibition signals and miscoding of direction in dorsomedial striatum. Frontiers in integrative neuroscience *6*, 69.

Burke, K.A., Franz, T.M., Miller, D.N., and Schoenbaum, G. (2008). The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. Nature *454*, 340-344.

Burton, A.C., Kashtelyan, V., Bryden, D.W., and Roesch, M.R. (2013). Increased Firing to Cues That Predict Low-Value Reward in the Medial Orbitofrontal Cortex. Cerebral cortex (New York, NY : 1991).

Cacciapaglia, F., Saddoris, M.P., Wightman, R.M., and Carelli, R.M. (2012). Differential dopamine release dynamics in the nucleus accumbens core and shell track distinct aspects of goal-directed behavior for sucrose. Neuropharmacology *62*, 2050-2056.

Cacciapaglia, F., Wightman, R.M., and Carelli, R.M. (2011). Rapid dopamine signaling differentially modulates distinct microcircuits within the nucleus accumbens during sucrose-directed behavior. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 13860-13869.

Cai, X., Kim, S., and Lee, D. (2011). Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. Neuron *69*, 170-182.

Cameron, C., and Carelli, R.M. (2012). Cocaine abstinence alters nucleus accumbens firing dynamics during goal-directed behaviors for cocaine and sucrose. In The European journal of neuroscience, pp. 940-951.

Cardinal, R.N., Parkinson, J.A., Hall, J., and Everitt, B.J. (2002a). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. In Neuroscience and biobehavioral reviews, pp. 321-352.

Cardinal, R.N., Parkinson, J.A., Lachenal, G., Halkerston, K.M., Rudarakanchana, N., Hall, J., Morrison, C.H., Howes, S.R., Robbins, T.W., and Everitt, B.J. (2002b). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. Behav Neurosci *116*, 553-567.

Cardinal, R.N., Pennicott, D.R., Sugathapala, C.L., Robbins, T.W., and Everitt, B.J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. Science *292*, 2499-2501.

Carelli, R.M. (2002). Nucleus accumbens cell firing during goal-directed behaviors for cocaine vs. 'natural' reinforcement. Physiology & behavior *76*, 379-387.

Carelli, R.M. (2004). Nucleus accumbens cell firing and rapid dopamine signaling during goal-directed behaviors in rats. Neuropharmacology *47 Suppl 1*, 180-189.

Carelli, R.M., and Deadwyler, S.A. (1994). A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 7735-7746.

Carelli, R.M., and Ijames, S.G. (2001). Selective activation of accumbens neurons by cocaine-associated stimuli during a water/cocaine multiple schedule. Brain research *907*, 156-161.

Carelli, R.M., Ijames, S.G., and Crumling, A.J. (2000). Evidence that separate neural circuits in the nucleus accumbens encode cocaine versus "natural" (water and food) reward. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 4255-4266.

Carelli, R.M., and West, M.O. (1991). Representation of the body by single neurons in the dorsolateral striatum of the awake, unrestrained rat. In The Journal of comparative neurology, pp. 231-249.

Carelli, R.M., Wolske, M., and West, M.O. (1997). Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 1804-1814.

Carli, M., Jones, G.H., and Robbins, T.W. (1989). Effects of unilateral dorsal and ventral striatal dopamine depletion on visual neglect in the rat: a neural and behavioural analysis. In Neuroscience, pp. 309-327.

Chang, J.Y., Chen, L., Luo, F., Shi, L.H., and Woodward, D.J. (2002). Neuronal responses in the frontal cortico-basal ganglia system during delayed matching-to-sample task: ensemble recording in freely moving rats. Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale *142*, 67-80.

Chaudhri, N., Sahuque, L.L., Schairer, W.W., and Janak, P.H. (2010). Separable roles of the nucleus accumbens core and shell in context- and cue-induced alcohol-seeking. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *35*, 783-791.

Cheatwood, J.L., Reep, R.L., and Corwin, J.V. (2003). The associative striatum: cortical and thalamic projections to the dorsocentral striatum in rats. Brain research *968*, 1-14.

Cheer, J.F., Aragona, B.J., Heien, M.L., Seipel, A.T., Carelli, R.M., and Wightman, R.M. (2007). Coordinated accumbal dopamine release and neural activity drive goal-directed behavior. Neuron *54*, 237-244.

Chevalier, G., and Deniau, J.M. (1990). Disinhibition as a basic process in the expression of striatal functions. Trends Neurosci *13*, 277-280.

Chuhma, N., Tanaka, K.F., Hen, R., and Rayport, S. (2011). Functional connectome of the striatal medium spiny neuron. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 1183-1192.

Clark, J.J., Hollon, N.G., and Phillips, P.E. (2012). Pavlovian valuation systems in learning and decision making. Current opinion in neurobiology *22*, 1054-1061.

Coffield, J. A., K. K. Bowen and V. Miletic (1992). Retrograde tracing of projections between the nucleus submedius, the ventrolateral orbital cortex, and the midbrain in the rat. J Comp Neurol 321(3): 488-499.

Cook, D., and Kesner, R.P. (1988). Caudate nucleus and memory for egocentric localization. In Behav Neural Biol, pp. 332-343.

Corbit, L.H., and Balleine, B.W. (2011). The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 11786-11794.

Corbit, L.H., Muir, J.L., and Balleine, B.W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 3251-3260.

Cui, G., Jun, S.B., Jin, X., Pham, M.D., Vogel, S.S., Lovinger, D.M., and Costa, R.M. (2013). Concurrent activation of striatal direct and indirect pathways during action initiation. Nature *494*, 238-242.

Dalley, J. W., R. N. Cardinal and T. W. Robbins (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. Neurosci Biobehav Rev 28(7): 771-784.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron *69*, 1204-1215.

Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature neuroscience *8*, 1704-1711.

Day, J.J., Jones, J.L., and Carelli, R.M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. The European journal of neuroscience *33*, 308-321.

Day, J.J., Jones, J.L., Wightman, R.M., and Carelli, R.M. (2010). Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. Biological psychiatry *68*, 306-309.

Day, J.J., Roitman, M.F., Wightman, R.M., and Carelli, R.M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nature neuroscience *10*, 1020-1028.

Day, J.J., Wheeler, R.A., Roitman, M.F., and Carelli, R.M. (2006). Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. The European journal of neuroscience *23*, 1341-1351.

Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. Current opinion in neurobiology *18*, 185-196.

de Araujo, I.E., Gutierrez, R., Oliveira-Maia, A.J., Pereira, A., Jr., Nicolelis, M.A., and Simon, S.A. (2006). Neural ensemble coding of satiety states. Neuron *51*, 483-494.

Delamater, A.R. (2007). The role of the orbitofrontal cortex in sensory-specific encoding of associations in pavlovian and instrumental conditioning. Annals of the New York Academy of Sciences *1121*, 152-173.

DeLong, M.R. (1990). Primate models of movement disorders of basal ganglia origin. Trends Neurosci *13*, 281-285.

Deniau, J.M., and Chevalier, G. (1985). Disinhibition as a basic process in the expression of striatal functions. II. The striato-nigral influence on thalamocortical cells of the ventromedial thalamic nucleus. Brain research *334*, 227-233.

Deniau, J.M., Menetrey, A., and Charpier, S. (1996). The lamellar organization of the rat substantia nigra pars reticulata: segregated patterns of striatal afferents and relationship to the topography of corticostriatal projections. Neuroscience *73*, 761-781.

Deniau, J.M., Menetrey, A., and Thierry, A.M. (1994). Indirect nucleus accumbens input to the prefrontal cortex via the substantia nigra pars reticulata: a combined anatomical and electrophysiological study in the rat. Neuroscience *61*, 533-545.

Devan, B.D., Hong, N.S., and Mcdonald, R. (2011). Parallel associative processing in the dorsal striatum: segregation of stimulus-response and cognitive control subregions. In Neurobiology of learning and memory, pp. 95-120.

Di Ciano, P., and Everitt, B.J. (2001). Dissociable effects of antagonism of NMDA and AMPA/KA receptors in the nucleus accumbens core and shell on cocaine-seeking behavior. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology *25*, 341-360.

Ding, D.C., Gabbott, P.L., and Totterdell, S. (2001). Differences in the laminar origin of projections from the medial prefrontal cortex to the nucleus accumbens shell and core regions in the rat. Brain research *917*, 81-89.

Döbrössy, M.D., and Dunnett, S.B. (1997). Unilateral striatal lesions impair response execution on a lateralised choice reaction time task. In Behav Brain Res, pp. 159-171.

Drevets, W.C. (2000). Functional anatomical abnormalities in limbic and prefrontal cortical structures in major depression. Progress in brain research *126*, 413-431.

Eagle, D.M., Baunez, C., Hutcheson, D.M., Lehmann, O., Shah, A.P., and Robbins, T.W. (2008). Stop-signal reaction-time task performance: role of prefrontal cortex and subthalamic nucleus. Cerebral cortex (New York, NY : 1991) *18*, 178-188.

Estes, W.K. (1948). Discriminative conditioning; effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. Journal of experimental psychology *38*, 173-177.

Fanelli, R., Klein, J., Reese, R., and Robinson, D. (2013). Dorsomedial and dorsolateral striatum exhibit distinct phasic neuronal activity during alcohol self-administration in rats. In The European journal of neuroscience, pp. 2637-2648.

Featherstone, R.E., and McDonald, R.J. (2004a). Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a stimulus-response-based instrumental discrimination task, while sparing conditioned place preference learning. Neuroscience *124*, 23-31.

Featherstone, R.E., and McDonald, R.J. (2004b). Dorsal striatum and stimulus–response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a simple discrimination task. Behavioural Brain Research *150*, 15-23.

Feierstein, C.E., Quirk, M.C., Uchida, N., Sosulski, D.L., and Mainen, Z.F. (2006). Representation of spatial goals in rat orbitofrontal cortex. Neuron *51*, 495-507.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. Science *299*, 1898-1902.

Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., Akers, C.A., Clinton, S.M., Phillips, P.E., and Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. Nature *469*, 53-57.

Floresco, S.B., McLaughlin, R., and Haluk, D. (2008). Opposing roles for the nucleus accumbens core and shell in cue-induced reinstatement of food seeking behavior. In Neuroscience, pp. 877-884.

Frank, M.J., Seeberger, L.C., and O'Reilly R, C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science *306*, 1940-1943.

Fuchs, R.A., Evans, K.A., Parker, M.C., and See, R.E. (2004). Differential involvement of the core and shell subregions of the nucleus accumbens in conditioned cue-induced reinstatement of cocaine seeking in rats. Psychopharmacology *176*, 459-465.

Furuyashiki, T., Holland, P.C., and Gallagher, M. (2008). Rat orbitofrontal cortex separately encodes response and outcome information during performance of goal-directed behavior. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 5127-5138.

Gage, G.J., Stoetzner, C.R., Wiltschko, A.B., and Berke, J.D. (2010). Selective activation of striatal fast-spiking interneurons during choice execution. In Neuron, pp. 466-479.

Gallagher, M., McMahan, R.W., and Schoenbaum, G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. The Journal of neuroscience : the official journal of the Society for Neuroscience *19*, 6610-6614.

Gerfen, C. R., T. M. Engber, L. C. Mahan, Z. Susel, T. N. Chase, F. J. Monsma, Jr. and D. R. Sibley (1990). "D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons." Science **250**(4986): 1429-1432.

Gerfen, C. R. and W. S. Young, 3rd (1988). "Distribution of striatonigral and striatopallidal peptidergic neurons in both patch and matrix compartments: an in situ hybridization histochemistry and fluorescent retrograde tracing study." Brain Res **460**(1): 161-167.

Gittis, A.H., Leventhal, D.K., Fensterheim, B.A., Pettibone, J.R., Berke, J.D., and Kreitzer, A.C. (2011). Selective inhibition of striatal fast-spiking interneurons causes dyskinesias. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 15727-15731.

Gittis, A.H., Nelson, A.B., Thwin, M.T., Palop, J.J., and Kreitzer, A.C. (2010). Distinct roles of GABAergic interneurons in the regulation of striatal output pathways. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 2223-2234.

Glascher, J., Daw, N., Dayan, P., and O'Doherty, J.P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron *66*, 585-595.

Goldstein, B.L., Barnett, B.R., Vasquez, G., Tobia, S.C., Kashtelyan, V., Burton, A.C., Bryden, D.W., and Roesch, M.R. (2012). Ventral striatum encodes past and predicted value independent of motor contingencies. The Journal of neuroscience : the official journal of the Society for Neuroscience *32*, 2027-2036.

Gorny, J.H., Gorny, B., Wallace, D.G., and Whishaw, I.Q. (2002). Fimbria-fornix lesions disrupt the dead reckoning (homing) component of exploratory behavior in mice. Learning & memory *9*, 387-394.

Graybiel, A.M. (2000). The basal ganglia. Current biology : CB *10*, R509-511.

Grillner, S., Hellgren, J., Menard, A., Saitoh, K., and Wikstrom, M.A. (2005). Mechanisms for selection of basic motor programs--roles for the striatum and pallidum. Trends Neurosci *28*, 364-370.

Groenewegen, H.J. (2003). The basal ganglia and motor control. Neural plasticity *10*, 107-120.

Groenewegen, H.J., Berendse, H.W., and Haber, S.N. (1993). Organization of the output of the ventral striatopallidal system in the rat: ventral pallidal efferents. Neuroscience *57*, 113-142.

Groenewegen, H.J., Berendse, H.W., Wolters, J.G., and Lohman, A.H. (1990). The anatomical relationship of the prefrontal cortex with the striatopallidal system, the thalamus and the amygdala: evidence for a parallel organization. Progress in brain research *85*, 95-116; discussion 116-118.

Groenewegen, H.J., and Russchen, F.T. (1984). Organization of the efferent projections of the nucleus accumbens to pallidal, hypothalamic, and mesencephalic structures: a tracing and immunohistochemical study in the cat. The Journal of comparative neurology *223*, 347-367.

Groenewegen, H.J., Vermeulen-Van der Zee, E., te Kortschot, A., and Witter, M.P. (1987). Organization of the projections from the subiculum to the ventral striatum in the rat. A study using anterograde transport of Phaseolus vulgaris leucoagglutinin. Neuroscience *23*, 103-120.

Groenewegen, H.J., Wright, C.I., Beijer, A.V., and Voorn, P. (1999). Convergence and segregation of ventral striatal inputs and outputs. Annals of the New York Academy of Sciences *877*, 49-63.

Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. Journal of Chemical Neuroanatomy *26*, 317-330.

Haber, S.N., and Brucker, J.L. (2009). Cognitive and limbic circuits that are affected by deep brain stimulation. Frontiers in bioscience (Landmark edition) *14*, 1823-1834.

Haber, S.N., Fudge, J.L., and McFarland, N.R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. The Journal of neuroscience : the official journal of the Society for Neuroscience *20*, 2369-2382.

Hakan, R.L., Berg, G.I., and Henriksen, S.J. (1992). Electrophysiological evidence for reciprocal connectivity between the nucleus accumbens septi and ventral pallidal region. Brain research *581*, 344-350.

Hassani, O.K., Cromwell, H.C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. In Journal of neurophysiology, pp. 2477-2489.

Heimer, L., Zahm, D.S., Churchill, L., Kalivas, P.W., and Wohltmann, C. (1991). Specificity in the projection patterns of accumbal core and shell in the rat. Neuroscience *41*, 89-125.

Hikosaka, K., and Watanabe, M. (2000). Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. In Cerebral cortex (New York, NY : 1991), pp. 263-271.

Hikosaka, O., Nakahara, H., Rand, M.K., Sakai, K., Lu, X., Nakamura, K., Miyachi, S., and Doya, K. (1999). Parallel neural networks for learning sequential procedures. Trends Neurosci *22*, 464-471.

Hikosaka, O., Sakamoto, M., and Usui, S. (1989a). Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. Journal of neurophysiology *61*, 780-798.

Hikosaka, O., Sakamoto, M., and Usui, S. (1989b). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. Journal of neurophysiology *61*, 814-832.

Hollerman, J.R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. In Journal of neurophysiology, pp. 947-963.

Hong, S., and Hikosaka, O. (2008). The globus pallidus sends reward-related signals to the lateral habenula. Neuron *60*, 720-729.

Hoover, W.B., and Vertes, R.P. (2011). Projections of the medial orbital and ventral orbital cortex in the rat. The Journal of comparative neurology *519*, 3766-3801.

Hori, Y., Minamimoto, T., and Kimura, M. (2009). Neuronal encoding of reward value and direction of actions in the primate putamen. Journal of neurophysiology *102*, 3530-3543.

Humphries, M., and Prescott, T. (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. In Progress in neurobiology, pp. 385-417.

Ikemoto, S. (2007). Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. Brain research reviews *56*, 27-78.

Ito, M., and Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 9861-9874.

Ito, R., Dalley, J.W., Howes, S.R., Robbins, T.W., and Everitt, B.J. (2000). Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. The Journal of neuroscience : the official journal of the Society for Neuroscience *20*, 7489-7495.

Izquierdo, A., and Jentsch, J.D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. Psychopharmacology *219*, 607-620.

Izquierdo, A., Suda, R.K., and Murray, E.A. (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. The Journal of neuroscience : the official journal of the Society for Neuroscience *24*, 7540-7548.

Jaeger, D., Kita, H., and Wilson, C.J. (1994). Surround inhibition among projection neurons is weak or nonexistent in the rat neostriatum. Journal of neurophysiology *72*, 2555-2558.

Janak, P.H., Chen, M.T., and Caulder, T. (2004). Dynamics of neural coding in the accumbens during extinction and reinstatement of rewarded behavior. Behav Brain Res *154*, 125-135.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. In Neural networks : the official journal of the International Neural Network Society, pp. 535-547.

Joel, D., and Weiner, I. (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. Neuroscience *63*, 363-379.

Joel, D., and Weiner, I. (1997). The connections of the primate subthalamic nucleus: indirect pathways and the open-interconnected scheme of basal ganglia-thalamocortical circuitry. Brain research Brain research reviews *23*, 62-78.

Joel, D., and Weiner, I. (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. In Neuroscience, pp. 451-474.

Jog, M.S. (1999). Building Neural Representations of Habits. Science *286*, 1745-1749.

Jones, J.L., Day, J.J., Wheeler, R.A., and Carelli, R.M. (2010). The basolateral amygdala differentially regulates conditioned neural responses within the nucleus accumbens core and shell. Neuroscience *169*, 1186-1198.

Jongen-Relo, A.L., Voorn, P., and Groenewegen, H.J. (1994). Immunohistochemical characterization of the shell and core territories of the nucleus accumbens in the rat. The European journal of neuroscience *6*, 1255-1264.

Kawaguchi, Y. (1993). Physiological, morphological, and histochemical characterization of three classes of interneurons in rat neostriatum. The Journal of neuroscience : the official journal of the Society for Neuroscience *13*, 4908-4923.

Kawaguchi, Y. (1995). Physiological subgroups of nonpyramidal cells with specific morphological characteristics in layer II/III of rat frontal cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *15*, 2638-2655.

Kawaguchi, Y., Wilson, C.J., and Emson, P.C. (1989). Intracellular recording of identified neostriatal patch and matrix spiny cells in a slice preparation preserving cortical inputs. Journal of neurophysiology *62*, 1052-1068.

Kelley, A.E. (1999). Functional specificity of ventral striatal compartments in appetitive behaviors. Annals of the New York Academy of Sciences *877*, 71-90.

Kelley, A.E., Baldo, B.A., Pratt, W.E., and Will, M.J. (2005). Corticostriatal-hypothalamic circuitry and food motivation: integration of energy, action and reward. Physiology & behavior *86*, 773-795.

Kelley, A.E., and Domesick, V.B. (1982). The distribution of the projection from the hippocampal formation to the nucleus accumbens in the rat: an anterograde- and retrograde-horseradish peroxidase study. Neuroscience *7*, 2321-2335.

Kelley, A.E., and Swanson, C.J. (1997). Feeding induced by blockade of AMPA and kainate receptors within the ventral striatum: a microinfusion mapping study. Behav Brain Res *89*, 107-113.

Kennedy, P.J., and Shapiro, M.L. (2009). Motivational states activate distinct hippocampal representations to guide goal-directed behaviors. Proceedings of the National Academy of Sciences of the United States of America *106*, 10805-10810.

Kennerley, S.W., Behrens, T.E., and Wallis, J.D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. In Nature neuroscience, pp. 1581-1589.

Kennerley, S.W., and Wallis, J.D. (2009a). Encoding of reward and space during a working memory task in the orbitofrontal cortex and anterior cingulate sulcus. Journal of neurophysiology *102*, 3352-3364.

Kennerley, S.W., and Wallis, J.D. (2009b). Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. The European journal of neuroscience *29*, 2061-2073.

Kepecs, A., Uchida, N., Zariwala, H.A., and Mainen, Z.F. (2008). Neural correlates, computation and behavioural impact of decision confidence. Nature *455*, 227-231.

Kerr, J.N., and Wickens, J.R. (2001). Dopamine D-1/D-5 receptor activation is required for long-term potentiation in the rat neostriatum in vitro. Journal of neurophysiology *85*, 117-124.

Kim, H., Lee, D., and Jung, M.W. (2013). Signals for previous goal choice persist in the dorsomedial, but not dorsolateral striatum of rats. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 52-63.

Kim, H., Sul, J., Huh, N., Lee, D., and Jung, M.W. (2009). Role of striatum in updating values of chosen actions. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 14701-14712.

Kim, Y.B., Huh, N., Lee, H., Baeg, E.H., Lee, D., and Jung, M.W. (2007). Encoding of action history in the rat ventral striatum. Journal of neurophysiology *98*, 3548-3556.

Kimchi, E.Y., and Laubach, M. (2009a). The dorsomedial striatum reflects response bias during learning. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 14891-14902.

Kimchi, E.Y., and Laubach, M. (2009b). Dynamic encoding of action selection by the medial striatum. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 3148-3159.

Kimchi, E.Y., Torregrossa, M.M., Taylor, J.R., and Laubach, M. (2009). Neuronal correlates of instrumental learning in the dorsal striatum. Journal of neurophysiology *102*, 475-489.

Kita, H. (1993). GABAergic circuits of the striatum. Progress in brain research *99*, 51-72.

Kita, H., Kosaka, T., and Heizmann, C.W. (1990). Parvalbumin-immunoreactive neurons in the rat neostriatum: a light and electron microscopic study. Brain research *536*, 1-15.

Knowlton, B.J., Mangels, J.A., and Squire, L.R. (1996). A neostriatal habit learning system in humans. Science *273*, 1399-1402.

Kobayashi, S., Pinto de Carvalho, O., and Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 534-544.

Koós, T., and Tepper, J.M. (1999). Inhibitory control of neostriatal projection neurons by GABAergic interneurons. In Nature neuroscience, pp. 467-472.

Koos, T., Tepper, J.M., and Wilson, C.J. (2004). Comparison of IPSCs evoked by spiny and fast-spiking neurons in the neostriatum. The Journal of neuroscience : the official journal of the Society for Neuroscience *24*, 7916-7922.

Krause, M., German, P.W., Taha, S.A., and Fields, H.L. (2010). A pause in nucleus accumbens neuron firing is required to initiate and maintain feeding. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 4746-4756.

Kravitz, A.V., Freeze, B.S., Parker, P.R., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature *466*, 622-626.

Kravitz, A.V., Tye, L.D., and Kreitzer, A.C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. Nature neuroscience *15*, 816-818.

Kubota, Y., Liu, J., Hu, D., DeCoteau, W.E., Eden, U.T., Smith, A.C., and Graybiel, A.M. (2009). Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. Journal of neurophysiology *102*, 2142-2160.

Lau, B., and Glimcher, P.W. (2007). Action and outcome encoding in the primate caudate nucleus. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 14502-14514.

Lau, B., and Glimcher, P.W. (2008). Value representations in the primate striatum during matching behavior. Neuron *58*, 451-463.

Lavoie, A.M., and Mizumori, S.J. (1994). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of rats. In Brain research, pp. 157-168.

Levy, D.J., and Glimcher, P.W. (2011). Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 14693-14707.

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. Journal of neurophysiology *67*, 145-163.

Lovibond, P.F. (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. Journal of experimental psychology Animal behavior processes *9*, 225-247.

Lu, X.Y., Ghasemzadeh, M.B., and Kalivas, P.W. (1998). Expression of D1 receptor, D2 receptor, substance P and enkephalin messenger RNAs in the neurons projecting from the nucleus accumbens. Neuroscience *82*, 767-780.

Luk, C.H., and Wallis, J.D. (2013). Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. The Journal of neuroscience : the official journal of the Society for Neuroscience *33*, 1864-1871.

Mallet, N., Le Moine, C., Charpier, S., and Gonon, F. (2005). Feedforward inhibition of projection neurons by fast-spiking GABA interneurons in the rat striatum in vivo. The Journal of neuroscience : the official journal of the Society for Neuroscience *25*, 3857-3869.

Mar, A.C., Walker, A.L., Theobald, D.E., Eagle, D.M., and Robbins, T.W. (2011). Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 6398-6404.

Matamales, M., Bertran-Gonzalez, J., Salomon, L., Degos, B., Deniau, J.M., Valjent, E., Herve, D., and Girault, J.A. (2009). Striatal medium-sized spiny neurons: identification by nuclear staining and study of neuronal subpopulations in BAC transgenic mice. PloS one *4*, e4770.

Maurice, N., Deniau, J.M., Menetrey, A., Glowinski, J., and Thierry, A.M. (1997). Position of the ventral pallidum in the rat prefrontal cortex-basal ganglia circuit. Neuroscience *80*, 523-534.

Maurin, Y., Banrezes, B., Menetrey, A., Mailly, P., and Deniau, J.M. (1999). Three-dimensional distribution of nigrostriatal neurons in the rat: relation to the topography of striatonigral projections. Neuroscience *91*, 891-909.

McDannald, M.A., Lucantonio, F., Burke, K.A., Niv, Y., and Schoenbaum, G. (2011). Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 2700-2705.

McDannald, M.A., Takahashi, Y.K., Lopatina, N., Pietras, B.W., Jones, J.L., and Schoenbaum, G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. The European journal of neuroscience *35*, 991-996.

McDonald, R.J., and Hong, N.S. (2004). A dissociation of dorso-lateral striatum and amygdala function on the same stimulus-response habit task. Neuroscience *124*, 507-513.

McDonald, R.J., and White, N.M. (1993). A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. Behav Neurosci *107*, 3-22.

McGinty, V.B., Lardeux, S., Taha, S.A., Kim, J.J., and Nicola, S.M. (2013). Invigoration of reward seeking by cue and proximity encoding in the nucleus accumbens. Neuron *78*, 910-922.

Minamimoto, T., Hori, Y., and Kimura, M. (2005). Complementary process to response bias in the centromedian nucleus of the thalamus. Science *308*, 1798-1801.

Minamimoto, T., Hori, Y., and Kimura, M. (2009). Roles of the thalamic CM-PF complex-Basal ganglia circuit in externally driven rebias of action. Brain research bulletin *78*, 75-79.

Mink, J.W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. In Progress in neurobiology, pp. 381-425.

Mizumori, S.J., Puryear, C.B., and Martig, A.K. (2009). Basal ganglia contributions to adaptive navigation. Behav Brain Res *199*, 32-42.

Mizumori, S.J., Yeshenko, O., Gill, K.M., and Davis, D.M. (2004). Parallel processing across neural systems: implications for a multiple memory system hypothesis. Neurobiology of learning and memory *82*, 278-298.

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

Mogenson, G.J., Swanson, L.W., and Wu, M. (1983). Neural projections from nucleus accumbens to globus pallidus, substantia innominata, and lateral preoptic-lateral hypothalamic area: an anatomical and electrophysiological investigation in the rat. The Journal of neuroscience : the official journal of the Society for Neuroscience *3*, 189-202.

Montague, P.R. (1993). Transforming sensory experience into structural change. Proceedings of the National Academy of Sciences of the United States of America *90*, 6379-6380.

Montague, P.R. (1995). Integrating information at single synaptic connections. Proceedings of the National Academy of Sciences of the United States of America *92*, 2424-2425.

Montague, P.R., and Sejnowski, T.J. (1994). The predictive brain: temporal coincidence and temporal order in synaptic learning mechanisms. Learning & memory *1*, 1-33.

Morrison, S.E., and Salzman, C.D. (2009). The convergence of information about rewarding and aversive stimuli in single neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 11471-11483.

Nakamura, K., Santos, G.S., Matsuzaki, R., and Nakahara, H. (2012). Differential reward coding in the subdivisions of the primate caudate during an oculomotor task. The Journal of neuroscience : the official journal of the Society for Neuroscience *32*, 15963-15982.

Nicola, S.M. (2010). The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 16585-16600.

Nicola, S.M., Yun, I.A., Wakabayashi, K.T., and Fields, H.L. (2004a). Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task. In Journal of neurophysiology, pp. 1840-1865.

Nicola, S.M., Yun, I.A., Wakabayashi, K.T., and Fields, H.L. (2004b). Firing of nucleus accumbens neurons during the consummatory phase of a discriminative stimulus task depends on previous reward predictive cues. In Journal of neurophysiology, pp. 1866-1882.

Niv, Y., Joel, D., and Dayan, P. (2006). A normative perspective on motivation. Trends in cognitive sciences *10*, 375-381.

Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., and Rushworth, M.F. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. In Proceedings of the National Academy of Sciences of the United States of America.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science *304*, 452-454.

O'Doherty, J.P., Deichmann, R., Critchley, H.D., and Dolan, R.J. (2002). Neural responses during anticipation of a primary taste reward. Neuron *33*, 815-826.

Ongur, D. and J. L. Price (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb Cortex 10(3): 206-219.

Opris, I., Hampson, R.E., and Deadwyler, S.A. (2009). The encoding of cocaine vs. natural rewards in the striatum of nonhuman primates: categories with different activations. Neuroscience *163*, 40-54.

Ostlund, S.B., and Balleine, B.W. (2007). Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. The Journal of neuroscience : the official journal of the Society for Neuroscience *27*, 4819-4825.

Oyama, K., Hernadi, I., Iijima, T., and Tsutsui, K. (2010). Reward prediction error coding in dorsal striatal neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 11447-11457.

Packard, M.G., Hirsh, R., and White, N.M. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. The Journal of neuroscience : the official journal of the Society for Neuroscience *9*, 1465-1472.

Packard, M.G., and Knowlton, B.J. (2002). Learning and memory functions of the Basal Ganglia. Annual review of neuroscience *25*, 563-593.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. Annual review of neuroscience *34*, 333-359.

Padoa-Schioppa, C., and Assad, J.A. (2006). Neurons in the orbitofrontal cortex encode economic value. Nature *441*, 223-226.

Padoa-Schioppa, C., and Cai, X. (2011). The orbitofrontal cortex and the computation of subjective value: consolidated concepts and new perspectives. Annals of the New York Academy of Sciences *1239*, 130-137.

Parent, A., and Hazrati, L.N. (1995). Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. Brain research Brain research reviews *20*, 91-127.

Parkinson, J.A., Olmstead, M.C., Burns, L.H., Robbins, T.W., and Everitt, B.J. (1999). Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. The Journal of neuroscience : the official journal of the Society for Neuroscience *19*, 2401-2411.

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

Parthasarathy, H.B., and Graybiel, A.M. (1997). Cortically driven immediate-early gene expression reflects modular influence of sensorimotor cortex on identified striatal neurons in the squirrel monkey. The Journal of neuroscience : the official journal of the Society for Neuroscience *17*, 2477-2491.

Partridge, J.G., Tang, K.C., and Lovinger, D.M. (2000). Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. Journal of neurophysiology *84*, 1422-1429.

Pasquereau, B., Nadjar, A., Arkadir, D., Bezard, E., Goillandeau, M., Bioulac, B., Gross, C.E., and Boraud, T. (2007). Shaping of motor responses by incentive values through the basal ganglia. The Journal of neuroscience : the official journal of the Society for Neuroscience *27*, 1176-1183.

Pasupathy, A., and Miller, E.K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. In Nature, pp. 873-876.

Paxinos G, Watson C. (1997). The Rat Brain in Stereotaxic Coordinates, Compact 3rd ed. London: Academic Press. p 11–15.

Pennartz, C.M., Groenewegen, H.J., and Lopes da Silva, F.H. (1994). The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomical data. Progress in neurobiology *42*, 719-761.

Pickens, C.L., Saddoris, M.P., Gallagher, M., and Holland, P.C. (2005). Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. Behav Neurosci *119*, 317-322.

Pickens, C.L., Saddoris, M.P., Setlow, B., Gallagher, M., Holland, P.C., and Schoenbaum, G. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. In Journal of Neuroscience, pp. 11078-11084.

Planert, H., Szydlowski, S.N., Hjorth, J.J., Grillner, S., and Silberberg, G. (2010). Dynamics of synaptic transmission between fast-spiking interneurons and striatal projection neurons of the direct and indirect pathways. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 3499-3507.

Price, J.L. (1985). Beyond the primary olfactory cortex: olfactory-related areas in the neocortex, thalamus, and hypothalamus. Chemical Senses 10: 235–258.

Price, J. L., S. T. Carmichael and W. C. Drevets (1996). Networks related to the orbital and medial prefrontal cortex; a substrate for emotional behavior? Prog Brain Res 107: 523-536.

Price, J. L. (2007). Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. Ann N Y Acad Sci 1121: 54-71.

Ragozzino, M.E., Ragozzino, K.E., Mizumori, S.J., and Kesner, R.P. (2002). Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. In Behav Neurosci, pp. 105-115.

Rangel, A., Camerer, C., and Montague, P.R. (2008). A framework for studying the neurobiology of value-based decision making. Nature reviews Neuroscience *9*, 545-556.

Reading, P.J., Dunnett, S.B., and Robbins, T.W. (1991). Dissociable roles of the ventral, medial and lateral striatum on the acquisition and performance of a complex visual stimulus-response habit. Behav Brain Res *45*, 147-161.

Redgrave, P., and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? Nature reviews Neuroscience *7*, 967-975.

Redgrave, P., Prescott, T.J., and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? In Neuroscience, pp. 1009-1023.

Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M.C., Lehericy, S., Bergman, H., Agid, Y., DeLong, M.R., and Obeso, J.A. (2010). Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. Nature reviews Neuroscience *11*, 760-772.

Peter Redgrave (2007), Scholarpedia, 2(6):1825.

Redish, A.D., Jensen, S., and Johnson, A. (2008). A unified framework for addiction: vulnerabilities in the decision process. The Behavioral and brain sciences *31*, 415-437; discussion 437-487.

Reep, R. L., H. C. Chandler, V. King and J. V. Corwin (1994). Rat posterior parietal cortex: topography of corticocortical and thalamic connections. Exp Brain Res 100(1): 67-84.

Reep, R. L., J. V. Corwin and V. King (1996). Neuronal connections of orbital cortex in rats: topography of cortical and thalamic afferents. Exp Brain Res 111(2): 215-232.

Robinson, D.L., and Carelli, R.M. (2008). Distinct subsets of nucleus accumbens neurons encode operant responding for ethanol versus water. The European journal of neuroscience *28*, 1887-1894.

Roesch, M.R., and Olson, C.R. (2004). Neuronal activity related to reward value and motivation in primate frontal cortex. Science *304*, 307-310.

Roesch, M.R., and Olson, C.R. (2007). Neuronal activity related to anticipated reward in frontal cortex: does it represent value or reflect motivation? Annals of the New York Academy of Sciences *1121*, 431-446.

Roesch, M.R., Singh, T., Brown, P.L., Mullins, S.E., and Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between

differently delayed or sized rewards. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 13365-13376.

Roesch, M.R., Taylor, A.R., and Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. Neuron *51*, 509-520.

Roitman, M.F., Stuber, G.D., Phillips, P.E., Wightman, R.M., and Carelli, R.M. (2004). Dopamine operates as a subsecond modulator of food seeking. The Journal of neuroscience : the official journal of the Society for Neuroscience *24*, 1265-1271.

Rolls, E.T. (1989). Information processing in the taste system of primates. The Journal of experimental biology *146*, 141-164.

Rolls, E.T. (2000). The orbitofrontal cortex and reward. Cerebral cortex (New York, NY : 1991) *10*, 284-294.

Rolls, E.T. (2004). The functions of the orbitofrontal cortex. Brain and cognition *55*, 11-29.

Rolls, E.T., and Baylis, L.L. (1994). Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *14*, 5437-5452.

Rolls, E.T., Critchley, H.D., Mason, R., and Wakeman, E.A. (1996). Orbitofrontal cortex neurons: role in olfactory and visual association learning. In Journal of neurophysiology, pp. 1970-1981.

Rolls, E.T., Thorpe, S.J., and Maddison, S.P. (1983). Responses of striatal neurons in the behaving monkey. 1. Head of the caudate nucleus. Behav Brain Res *7*, 179-210.

Romo, R., and Schultz, W. (1990). Dopamine neurons of the monkey midbrain: contingencies of responses to active touch during self-initiated arm movements. Journal of neurophysiology *63*, 592-606.

Roop, R.G., Hollander, J.A., and Carelli, R.M. (2002). Accumbens activity during a multiple schedule for water and sucrose reinforcement in rats. Synapse (New York, NY) *43*, 223-226.

Root, D.H., Tang, C.C., Ma, S., Pawlak, A., and West, M. (2010). Absence of cue-evoked firing in rat dorsolateral striatum neurons. In Behav Brain Res, pp. 23-32.

Rothemund, Y., Preuschhof, C., Bohner, G., Bauknecht, H.C., Klingebiel, R., Flor, H., and Klapp, B.F. (2007). Differential activation of the dorsal striatum by high-calorie visual food stimuli in obese individuals. NeuroImage *37*, 410-421.

Rudebeck, P.H., and Murray, E.A. (2011a). Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. Annals of the New York Academy of Sciences *1239*, 1-13.

Rudebeck, P.H., and Murray, E.A. (2011b). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 10569-10578.

Rudebeck, P.H., Walton, M.E., Smyth, A.N., Bannerman, D.M., and Rushworth, M.F. (2006). Separate neural pathways process different decision costs. Nature neuroscience *9*, 1161-1168.

Samejima, K., and Doya, K. (2007). Multiple representations of belief states and action values in corticobasal ganglia loops. Annals of the New York Academy of Sciences *1104*, 213-228.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. Science *310*, 1337-1340.

Sargolini, F., Florian, C., Oliverio, A., Mele, A., and Roullet, P. (2003). Differential involvement of NMDA and AMPA receptors within the nucleus accumbens in consolidation of information necessary for place navigation and guidance strategy of mice. Learning & memory *10*, 285-292.

Scarlet, J., Delamater, A.R., Campese, V., Fein, M., and Wheeler, D.S. (2012). Differential involvement of the basolateral amygdala and orbitofrontal cortex in the formation of sensory-specific associations in conditioned flavor preference and magazine approach paradigms. The European journal of neuroscience *35*, 1799-1809.

Schmidt, R., Leventhal, D.K., Mallet, N., Chen, F., and Berke, J.D. (2013). Canceling actions involves a race between basal ganglia pathways. Nature neuroscience *16*, 1118-1124.

Schmitzer-Torbert, N.C., and Redish, A.D. (2004). Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. In Journal of neurophysiology, pp. 2259-2272.

Schmitzer-Torbert, N.C., and Redish, A.D. (2008). Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. In Neuroscience, pp. 349-360.

Schoenbaum, G., Chiba, A.A., and Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. In Nature neuroscience, pp. 155-159.

Schoenbaum, G., Nugent, S.L., Saddoris, M.P., and Setlow, B. (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport *13*, 885-890.

Schoenbaum, G. and G. R. Esber (2010). How do you (estimate you will) like them apples? Integration as a defining trait of orbitofrontal function. Curr Opin Neurobiol 20(2): 205-211.

Schoenbaum, G., and Roesch, M. (2005). Orbitofrontal cortex, associative learning, and expectancies. Neuron *47*, 633-636.

Schoenbaum, G., Roesch, M.R., Stalnaker, T.A., and Takahashi, Y.K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. In Nature reviews Neuroscience, pp. 885-892.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. Annual review of psychology *57*, 87-115.

Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. The Journal of neuroscience : the official journal of the Society for Neuroscience *13*, 900-913.

Schultz, W., Tremblay, L., and Hollerman, J.R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. Cerebral cortex (New York, NY : 1991) *10*, 272-284.

Seo, H., and Lee, D. (2009). Behavioral and neural changes after gains and losses of conditioned reinforcers. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 3627-3641.

Seo, M., Lee, E., and Averbeck, B.B. (2012). Action selection and action value in frontal-striatal circuits. Neuron *74*, 947-960.

Sesack, S.R., Carr, D.B., Omelchenko, N., and Pinto, A. (2003). Anatomical substrates for glutamate-dopamine interactions: evidence for specificity of connections and extrasynaptic actions. Annals of the New York Academy of Sciences *1003*, 36-52.

Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning. In Neuron, pp. 625-636.

Shidara, M., and Richmond, B.J. (2004). Differential encoding of information about progress through multi-trial reward schedules by three groups of ventral striatal neurons. Neuroscience research *49*, 307-314.

Simmons, J.M., Ravel, S., Shidara, M., and Richmond, B.J. (2007). A comparison of reward-contingent neuronal activity in monkey orbitofrontal cortex and ventral striatum: guiding actions toward rewards. In Annals of the New York Academy of Sciences, pp. 376-394.

Simon, D.A., and Daw, N.D. (2011). Neural correlates of forward planning in a spatial decision task in humans. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 5526-5539.

Skinner, B. (1938) The Behavior of Organisms. Appleton-Century-Crofts, New York.

Smith, K.S., Tindell, A.J., Aldridge, J.W., and Berridge, K.C. (2009). Ventral pallidum roles in reward and motivation. In Behavioural Brain Research, pp. 155-167.

Soderpalm, A.H., and Berridge, K.C. (2000). Food intake after diazepam, morphine or muscimol: microinjections In the nucleus accumbens shell. Pharmacology, biochemistry, and behavior *66*, 429-434.

Stalnaker, T.A., Calhoon, G.G., Ogawa, M., Roesch, M.R., and Schoenbaum, G. (2010). Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Frontiers in integrative neuroscience *4*, 12.

Stalnaker, T.A., Calhoon, G.G., Ogawa, M., Roesch, M.R., and Schoenbaum, G. (2012). Reward prediction error signaling in posterior dorsomedial striatum is action specific. The Journal of neuroscience : the official journal of the Society for Neuroscience *32*, 10296-10305.

Stanton, G.B., Goldberg, M.E., and Bruce, C.J. (1988). Frontal eye field efferents in the macaque monkey: I. Subcortical pathways and topography of striatal and thalamic terminal fields. The Journal of comparative neurology *271*, 473-492.

Stratford, T.R., and Kelley, A.E. (1997). GABA in the nucleus accumbens shell participates in the central regulation of feeding behavior. The Journal of neuroscience : the official journal of the Society for Neuroscience *17*, 4434-4440.

Stratford, T.R., Swanson, C.J., and Kelley, A.E. (1998). Specific changes in food intake elicited by blockade or activation of glutamate receptors in the nucleus accumbens shell. In Behav Brain Res, pp. 43-50.

Stratford, T.R., and Wirtshafter, D. (2004). NPY mediates the feeding elicited by muscimol injections into the nucleus accumbens shell. Neuroreport *15*, 2673-2676.

Sul, J., Jo, S., Lee, D., and Jung, M.W. (2011). Role of rodent secondary motor cortex in value-based action selection. In Nature neuroscience, pp. 1202-1208.

Sul, J., Kim, H., Huh, N., Lee, D., and Jung, M.W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. In Neuron, pp. 449-460.

Sul, J.H., Jo, S., Lee, D., and Jung, M.W. (2011). Role of rodent secondary motor cortex in value-based action selection. Nature neuroscience *14*, 1202-1208.

Surmeier, D.J., Shen, W., Day, M., Gertler, T., Chan, S., Tian, X., and Plotkin, J.L. (2010). The role of dopamine in modulating the structure and function of striatal circuits. Progress in brain research *183*, 149-167.

Surmeier, D.J., Song, W.J., and Yan, Z. (1996). Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience *16*, 6579-6591.

Sutton, R.S., and Barto, A.G. (1998). Reinforce- ment Learning: An Introduction (Cambridge, MA: MIT Press).

Taha, S.A., and Fields, H.L. (2005). Encoding of palatability and appetitive behaviors by distinct neuronal populations in the nucleus accumbens. The Journal of neuroscience : the official journal of the Society for Neuroscience *25*, 1193-1202.

Taha, S.A., and Fields, H.L. (2006). Inhibitions of nucleus accumbens neurons encode a gating signal for reward-directed behavior. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 217-222.

Taha, S.A., Katsuura, Y., Noorvash, D., Seroussi, A., and Fields, H.L. (2009). Convergent, not serial, striatal and pallidal circuits regulate opioid-induced food intake. Neuroscience *161*, 718-733.

Taha, S.A., Nicola, S.M., and Fields, H.L. (2007). Cue-evoked encoding of movement planning and execution in the rat nucleus accumbens. The Journal of physiology *584*, 801-818.

Tai, L.H., Lee, A.M., Benavidez, N., Bonci, A., and Wilbrecht, L. (2012). Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. Nature neuroscience *15*, 1281-1289.

Takahashi, Y., Schoenbaum, G., and Niv, Y. (2008). Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model. Frontiers in neuroscience *2*, 86-99.

Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., and Schoenbaum, G. (2011). Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nature neuroscience *14*, 1590-1597.

Takikawa, Y., Kawagoe, R., and Hikosaka, O. (2004). A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. Journal of neurophysiology *92*, 2520-2529.

Tang, C.C., Pawlak, A., Prokopenko, V., and West, M. (2007). Changes in activity of the striatum during formation of a motor habit. In The European journal of neuroscience, pp. 1212-1227.

Taverna, S., Ilijic, E., and Surmeier, D.J. (2008). Recurrent collateral connections of striatal medium spiny neurons are disrupted in models of Parkinson's disease. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 5504-5512.

Tepper, J.M., and Bolam, J.P. (2004). Functional diversity and specificity of neostriatal interneurons. Current opinion in neurobiology *14*, 685-692.

Thorn, C.A., Atallah, H., Howe, M., and Graybiel, A.M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. Neuron *66*, 781-795.

Thorndike EL. 1911. *Animal Intelligence: Ex- perimental Studies*. New York: MacMillan

Thorpe, S.J., Rolls, E.T., and Maddison, S. (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. In Experimental brain research Experimentelle Hirnforschung Experimentation cerebrale, pp. 93-115.

Tindell, A.J., Berridge, K.C., and Aldridge, J.W. (2004). Ventral pallidal representation of pavlovian cues and reward: population and rate codes. The Journal of neuroscience : the official journal of the Society for Neuroscience *24*, 1058-1069.

Tindell, A.J., Smith, K.S., Pecina, S., Berridge, K.C., and Aldridge, J.W. (2006). Ventral pallidum firing codes hedonic reward: when a bad taste turns good. Journal of neurophysiology *96*, 2399-2409.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. Science *307*, 1642-1645.

Totterdell, S., and Meredith, G.E. (1997). Topographical organization of projections from the entorhinal cortex to the striatum of the rat. Neuroscience *78*, 715-729.

Tremblay, L., Hollerman, J.R., and Schultz, W. (1998). Modifications of reward expectation-related neuronal activity during learning in primate striatum. In Journal of neurophysiology, pp. 964-977.

Tremblay, L., and Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. In Nature, pp. 704-708.

Tremblay, L., and Schultz, W. (2000). Modifications of reward expectation-related neuronal activity during learning in primate orbitofrontal cortex. In Journal of neurophysiology, pp. 1877-1885.

Tricomi, E., Balleine, B.W., and O'Doherty, J.P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. The European journal of neuroscience *29*, 2225-2232.

Tricomi, E.M., Delgado, M.R., and Fiez, J.A. (2004). Modulation of caudate activity by action contingency. Neuron *41*, 281-292.

Tsujimoto, S., Genovesio, A., and Wise, S.P. (2009). Monkey orbitofrontal cortex encodes response choices near feedback time. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 2569-2574.

Usuda, I., Tanaka, K., and Chiba, T. (1998). Efferent projections of the nucleus accumbens in the rat with special reference to subdivision of the nucleus: biotinylated dextran amine study. Brain research *797*, 73-93.

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

Van Der Meer, M.A., and Redish, A.D. (2009a). Covert Expectation-of-Reward in Rat Ventral Striatum at Decision Points. In Frontiers in integrative neuroscience, pp. 1.

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

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

van der Meer, M.A., and Redish, A.D. (2011). Ventral striatum: a critical look at models of learning and evaluation. Current opinion in neurobiology *21*, 387-392.

van Duuren, E., Escamez, F.A., Joosten, R.N., Visser, R., Mulder, A.B., and Pennartz, C.M. (2007). Neural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor olfactory discrimination task. Learning & memory *14*, 446-456.

van Duuren, E., Lankelma, J., and Pennartz, C.M. (2008). Population coding of reward magnitude in the orbitofrontal cortex of the rat. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 8590-8603.

van Duuren, E., van der Plasse, G., Lankelma, J., Joosten, R.N., Feenstra, M.G., and Pennartz, C.M. (2009). Single-cell and population coding of expected reward probability in the orbitofrontal cortex of the rat. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 8965-8976.

van Groen, T., and Wyss, J.M. (1990). Extrinsic projections from area CA1 of the rat hippocampus: olfactory, cortical, subcortical, and bilateral hippocampal formation projections. The Journal of comparative neurology *302*, 515-528.

Vicente, A.F., Bermudez, M.A., Romero, M.d.C., Perez, R., and Gonzalez, F. (2012). Putamen neurons process both sensory and motor information during a complex task. In Brain research, pp. 70-81.

Voorn, P., Vanderschuren, L.J., Groenewegen, H.J., Robbins, T.W., and Pennartz, C.M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci *27*, 468-474.

Wall, N.R., De La Parra, M., Callaway, E.M., and Kreitzer, A.C. (2013). Differential innervation of direct- and indirect-pathway striatal projection neurons. Neuron *79*, 347-360.

Wallis, J.D. (2012). Cross-species studies of orbitofrontal cortex and value-based decision-making. Nature neuroscience *15*, 13-19.

Wallis, J.D., and Miller, E.K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. European Journal of Neuroscience *18*, 2069-2081.

Walton, M.E., Kennerley, S.W., Bannerman, D.M., Phillips, P.E., and Rushworth, M.F. (2006). Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. Neural networks : the official journal of the International Neural Network Society *19*, 1302-1314.

Whishaw, I.Q., Cassel, J.C., and Jarrad, L.E. (1995). Rats with fimbria-fornix lesions display a place response in a swimming pool: a dissociation between getting there and knowing where. The Journal of neuroscience : the official journal of the Society for Neuroscience *15*, 5779-5788.

White, N.M. (2009). Some highlights of research on the effects of caudate nucleus lesions over the past 200 years. Behav Brain Res *199*, 3-23.

Wiener, S.I. (1993). Spatial and behavioral correlates of striatal neurons in rats performing a self-initiated navigation task. In The Journal of neuroscience : the official journal of the Society for Neuroscience, pp. 3802-3817.

Wilkinson, L.S., Humby, T., Killcross, A.S., Torres, E.M., Everitt, B.J., and Robbins, T.W. (1998). Dissociations in dopamine release in medial prefrontal cortex and ventral striatum during the acquisition and extinction of classical aversive conditioning in the rat. The European journal of neuroscience *10*, 1019-1026.

Wilson, D.I., and Bowman, E.M. (2004). Nucleus accumbens neurons in the rat exhibit differential activity to conditioned reinforcers and primary reinforcers within a second-order schedule of saccharin reinforcement. The European journal of neuroscience *20*, 2777-2788.

Wiltschko, A.B., Gage, G.J., and Berke, J.D. (2008). Wavelet filtering before spike detection preserves waveform shape and enhances single-unit discrimination. Journal of neuroscience methods *173*, 34-40.

Wimmer, G.E., Daw, N.D., and Shohamy, D. (2012). Generalization of value in reinforcement learning by humans. The European journal of neuroscience *35*, 1092-1104.

Wise, R.A. (1978). Catecholamine theories of reward: a critical review. Brain research *152*, 215-247.

Wright, C.I., and Groenewegen, H.J. (1995). Patterns of convergence and segregation in the medial nucleus accumbens of the rat: relationships of prefrontal cortical, midline thalamic, and basal amygdaloid afferents. The Journal of comparative neurology *361*, 383-403.

Wunderlich, K., Dayan, P., and Dolan, R.J. (2012). Mapping value based planning and extensively trained choice in the human brain. Nature neuroscience *15*, 786-791.

Wunderlich, K., Rangel, A., and O'doherty, J. (2009). Neural computations underlying action-based decision making in the human brain. In Proceedings of the National Academy of Sciences of the United States of America, pp. 17199-17204.

Yamada, H., Inokawa, H., Matsumoto, M., Ueda, Y., and Kimura, M. (2011). Neuronal basis for evaluating selected action in the primate striatum. In The European journal of neuroscience, pp. 489-506.

Yamada, H., Matsumoto, M., and Kimura, M. (2004). Tonically active neurons in the primate caudate nucleus and putamen differentially encode instructed motivational outcomes of action. In Journal of Neuroscience, pp. 3500-3510.

Yamada, H., Matsumoto, N., and Kimura, M. (2007). History- and current instruction-based coding of forthcoming behavioral outcomes in the striatum. Journal of neurophysiology *98*, 3557-3567.

Yeshenko, O., Guazzelli, A., and Mizumori, S.J. (2004). Context-dependent reorganization of spatial and movement representations by simultaneously recorded

hippocampal and striatal neurons during performance of allocentric and egocentric tasks. Behav Neurosci *118*, 751-769.

Yin, H., and Knowlton, B. (2004). Contributions of striatal subregions to place and response learning. In Learning & memory, pp. 459-463.

Yin, H., and Knowlton, B. (2006). The role of the basal ganglia in habit formation. In Nature reviews Neuroscience, pp. 464-476.

Yin, H., Knowlton, B., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. In The European journal of neuroscience, pp. 181-189.

Yin, H., Knowlton, B., and Balleine, B.W. (2006). INActivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. In Behavioural Brain Research, pp. 189-196.

Yin, H., Ostlund, S.B., and Balleine, B.W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. In The European journal of neuroscience, pp. 1437-1448.

Yin, H., Ostlund, S.B., Knowlton, B., and Balleine, B.W. (2005a). The role of the dorsomedial striatum in instrumental conditioning. In The European journal of neuroscience, pp. 513-523.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2005b). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. The European journal of neuroscience *22*, 505-512.

Young, J.J., and Shapiro, M.L. (2011). Dynamic coding of goal-directed paths by orbital prefrontal cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *31*, 5989-6000.

Zaborszky, L., Alheid, G.F., Beinfeld, M.C., Eiden, L.E., Heimer, L., and Palkovits, M. (1985). Cholecystokinin innervation of the ventral striatum: a morphological and radioimmunological study. Neuroscience *14*, 427-453.

Zahm, D.S., and Brog, J.S. (1992). On the significance of subterritories in the "accumbens" part of the rat ventral striatum. Neuroscience *50*, 751-767.

Zastrow, A., Kaiser, S., Stippich, C., Walther, S., Herzog, W., Tchanturia, K., Belger, A., Weisbrod, M., Treasure, J., and Friederich, H.C. (2009). Neural correlates of impaired cognitive-behavioral flexibility in anorexia nervosa. The American journal of psychiatry *166*, 608-616.

Zhou, L., Furuta, T., and Kaneko, T. (2003). Chemical organization of projection neurons in the rat accumbens nucleus and olfactory tubercle. Neuroscience *120*, 783-798.