ABSTRACT

Title of Dissertation: THE LONG-TERM IMPACT OF PREVIOUS COCAINE SELF-ADMINISTRATION ON DECISION-MAKING AND STRIATAL CIRCUITRY

Amanda Claire Burton, Doctor of Philosophy, 2017

Dissertation directed by: Dr. Matthew Roesch, Associate Professor, Department of Psychology, Neuroscience and Cognitive Science Program

Current theories of addiction suggest that impaired decision-making observed in individuals that chronically abuse drugs reflects a decrease in goal-directed behaviors and an increase in habitual behaviors governed by neural representations of response-outcome and stimulus-response associations, respectively. The striatum is a critical input component to the basal ganglia, which is a complex set of subcortical brain structures involved in the selection and execution of actions. Striatal sub-regions are some of the first brain regions to be affected by drugs of abuse, yet we still do not fully understand how decision-making and neural correlates in these regions are affected by drug exposure or disruptions within the circuit. My project was designed to study behavioral and neural changes in the striatum after previous cocaine self-administration or pharmacological lesion while rats perform a complex decision-making task. I therefore implemented a cocaine self-administration or

pharmacological lesion protocol and recorded from single neurons in striatal sub-regions, specifically the nucleus accumbens core (NAc) and dorsal lateral striatum (DLS), during performance of an odor-guided decision-making task in which reward contingencies often changed. This task independently manipulated value of expected reward by changing the delay to or size of reward across a series of blocks of trials.

I found that previous cocaine self-administration made rats more impulsive, biasing choice behavior toward more immediate reward. After cocaine exposure, there were fewer task-responsive neurons in the NAc and in those that remained we observed diminished directional and value encoding compared to controls. Surprisingly, in the DLS I found evidence of increased response-outcome associations and no evidence of enhanced stimulus-response associations after cocaine exposure. After disrupting communication between the NAc and DLS, I found evidence of enhanced stimulus-response associations in the DLS during task performance. This suggests that cocaine exposure impacts decision-making and neural activity in the striatum that manifests in more complex ways than simply disrupting striatal circuitry as current theories of addiction suggest.

THE LONG-TERM IMPACT OF PREVIOUS COCAINE SELF-ADMINISTRATION ON DECISION-MAKING AND STRIATAL CIRCUITRY

by

Amanda Claire Burton

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Professor Matthew Roesch, Chair
Professor Donald Bolger, Dean's Representative
Professor Erica Glasper Andrews
Professor Elizabeth Quinlan
Professor Tracy Riggins

Dedication and Acknowledgements

I would like to dedicate this dissertation to the memory of my grandfather, Dr. Oswald James Burton and to my aunt, Professor and Dr. W. Anne Theibert. Thank you both for inspiring me to never stop learning and for your never-ending curiosity about my studies.

This dissertation would not exist without the help and support from many people. I would first like to thank my fiancé, Daniel Innocenti, for his unconditional support and love while helping me through the most stressful times in the last few years. I would also like to thank my parents, David and Debra Burton, for the many hours of dedicated help with writing, for moral and financial support, and for teaching me perseverance. I love all of you.

This dissertation would not be possible without multiple advisors. I would like to thank my first academic advisor, Professor Sue Adler and her husband, Professor Bill Patterson. Thank you for being an additional support network and always believing in me. Most of all, thank you to my dissertation advisor, Dr. Matthew Roesch. It has been an honor to work with you over the last few years, I have learned so much from you. Thank you for taking me on as an undergraduate many years ago.

An additional thank you to the members of the Roesch lab, for your help both inside and outside the lab, in alphabetical order: Dr. Gregory Bissonette, Dr. Daniel Bryden, Ronny Gentry, Vadim Kashtelyan, Brian Lee, Nina Lichtenberg, Stephen Tennyson, and Daniela Vazquez.

Table of Contents

Dedication and Acknowledgements ii
Table of Contents iii
List of Figures iv
Chapter 1: Introduction 1
General Introduction 1
A more detailed look within the reward circuit: the striatum 8
Synaptic plasticity in the striatum after cocaine exposure 10
Encoding in single neurons within striatum 13
The ventral striatum 14
The dorsal striatum 18
Addiction, decision-making, and how to investigate it in animal models 23
Odor-guided decision-making task: a modified version of a delay discounting task……..…….28
Chapter 2: Previous cocaine self-administration disrupts reward expectancy encoding in ventral striatum 34
Abstract 35
Introduction 36
Materials and Methods 38
Results 45
Discussion 61
Chapter 3: Prior cocaine self-administration increases response-outcome encoding that is divorced from actions selected in dorsal striatum 64
Abstract 65
Introduction 66
Materials and Methods 68
Results 72
Discussion 87
Chapter 4: Ventral striatum lesions enhance stimulus and response encoding in dorsal striatum 92
Abstract 93
Introduction 94
Materials and Methods 96
Results 99
Discussion 112
Chapter 5: General Discussion 114
Summary of Results 114
The impact of cocaine and lesions on behavior in the odor-guided decision-making task 121
Connecting the behavioral impact of cocaine and lesions to changes in neural signaling in the striatum 126
The impact of cocaine in the reward circuit and how our results fit within the circuit 129
Future directions 133
References 136

List of Figures

1.1 Reward circuitry and anatomy……....4
1.2 Task and behavior……..30
2.1 Task, behavior, and recording sites……....47
2.2 Previous cocaine-exposure diminishes value and directional encoding in NAc...51
2.3 Distribution of value indices for delay and size manipulations…….53
2.4 Firing on longer delays was attenuated in cocaine animals……...58
3.1 Task, behavior, and recording sites……73
3.2 Percentage of increasing and decreasing neurons in DLS modulated by task.....78
3.3 Single neuron examples of chosen-outcome and action-outcome encoding…….80
3.4 Action-outcome encoding is enhanced at the population level in cocaine-exposed rats……83
3.5 Neural selectivity emerges early in cocaine-exposed rats and was stimulus independent…….....86
4.1 Task, behavior, and recording/lesion sites……...……100
4.2 Stimulus and response encoding was enhanced after VS lesions…….....103
4.3 Stimulus and response encoding in decreasing- and increasing-type DS neurons was enhanced after VS lesions……...108
4.4 VS lesions caused temporary impairments of reward-guided decision-making..110

Chapter 1: Introduction

General Introduction

Drug addiction is a widespread disorder and a vital public health issue in the United States. In the most recent conducted drug use survey, more than 24 million Americans over age 12 had used illicit drugs in the past month. In addition, more than $700 billion is spent annually on the overall costs of substance abuse including healthcare, crime, and lost work productivity (NIDA Drug Facts, www.drugabuse.gov). While many advances have been made in recent decades in understanding how exposure to drugs of abuse affects the brain (Wise and Koob, 2014; Volkow and Morales, 2015; Everitt and Robbins, 2016), we still do not fully understand these effects and how they translate to common maladaptive behaviors seen after substance abuse. A more complete understanding of the neural mechanisms underlying disruptions in neural signaling and behavior after drug exposure is critical for improving treatment options for addicts and improving this significant health issue in the United States.

Decision-making is a key executive function that is central to normal everyday behavior that becomes disrupted during and after exposure to drugs of abuse (Hyman et al., 2006; Roesch et al., 2007b; Robbins et al., 2008; Stalnaker et al., 2009). In particular, the disruption of reward-based decision-making and its neural correlates are beginning to become untangled in the literature (Takahashi et al., 2007; Lucantonio et al., 2012; Burton et al., 2014a; Volkow and Morales, 2015). There are many brain regions that are involved in different aspects of reward-based learning and decision-making—learning about cues in the environment, encoding expected values

or outcomes related to these cues, modulating neural activity to generate motivation to specific cues related to valued outcomes, and combining these aspects to initiate actions to obtain the best and avoid the worst outcomes (Schultz, 2006; Balleine and O'Doherty, 2010; Burton et al., 2015; Bissonette and Roesch, 2016). Each of these is important to processing reward and responding accordingly in order to navigate and adapt to different environments. Although many brain regions are involved, a general reward-circuitry emerges when looking at the neural mechanisms underlying this type of decision-making that are essential to understanding the impact of drug exposure (Haber, Fudge, & McFarland, 2000; Haber & Knutson, 2009; Koob & Volkow, 2010; Stalnaker et al., 2009).

In brief, some of the brain regions making up this reward-circuit include the medial prefrontal cortex (mPFC), the orbitofrontal cortex (OFC), the basolateral amygdala (ABL), ventral regions of striatum—including the nucleus accumbens core (NAc) and nucleus accumbens shell (NAs), and dorsal regions of striatum—specifically dorsal medial striatum (DMS) and dorsal lateral striatum (DLS). There is evidence that early on in reward processing, OFC and ABL are important for animals to learn about valued reward options through midbrain dopaminergic modulation and encode expected or predicted outcomes, which is then passed along to multiple sub-regions of the basal ganglia, including the striatum, that control subsequent behaviors (Haber et al., 2000; Schultz, 2006; Haber and Knutson, 2009; Balleine and O'Doherty, 2010). After extensive learning, mPFC and DLS are critical for the formulation of habits, which do not require explicit representations of value.

The striatal regions—NAc, DMS, and DLS—work in parallel to incorporate information from upstream cortical regions and downstream regions to encode relevant information in the environment (Figure 1A). They help push the motor system towards optimal action selection to obtain valued outcomes and suppress the motor system to avoid aversive outcomes (Bromberg-Martin et al., 2010). More specifically, the NAc uses information about upcoming outcomes to modulate neural activity and motivate behavior to particular cues predicting these outcomes (Takahashi et al., 2007; Roesch et al., 2009; Goldstein et al., 2012; Bissonette et al., 2013; Burton et al., 2015). This information can then update more dorsal regions of striatum that govern action-related behaviors to obtain rewarded outcomes based on future environmental cues (Balleine and O'Doherty, 2010; Stalnaker et al., 2010; Burton et al., 2015). In addition to receiving information from the NAc, DMS and DLS are simultaneously receiving information from cortical regions and midbrain dopamine (DA) neurons to encode response-outcome (R-O) and stimulus-response (S-R) contingencies, respectively (explained in more detail below). R-O and S-R contingencies are incorporated with NAc updates about valued outcomes to propagate information forward to direct and indirect motor pathways responsible for generating or suppressing subsequent actions, respectively.

Phasic bursts of dopaminergic innervation from the midbrain provide reinforcing teaching signals to each of these regions in a reward prediction error fashion, in which the DA response is equivalent to reward occurred minus reward predicted (Schultz, 2006). There can be positive, negative, or no prediction error based on the result between what rewards occurred and what were expected (more

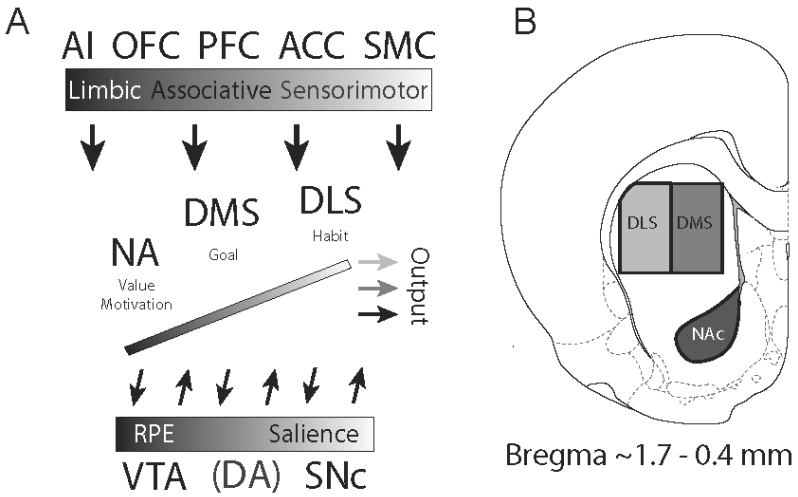

Figure 1: A. Both NAc and DMS receive projections from medial prefrontal cortex (mPFC) and anterior cingulate cortex (ACC), however NAc is more heavily innervated by agranular insular cortex (AI). DLS receives predominantly sensorimotor-related information from sensorimotor cortex (SMC), as opposed to more ventromedial parts of striatum, which receives visceral-related information (AI). Striatal areas in between NAc and DLS receive higher order associative information from mPFC and ACC. Functional differentiation can be recognized not only through cortical inputs, but also through afferents arising from amygdala, hippocampus and thalamus (Voorn et al., 2004). Amygdalostriatal projections are heaviest ventrally and progressively taper off in a dorsolateral direction, with sensorimotor parts of striatum being only very sparsely innervated. The medial (A8) and lateral (A10) dopamine neurons project predominantly dorsolaterally and ventromedially in striatum, respectively, whereas dopamine neurons more centrally located (A9) project broadly to the intermediary striatal zone, with some dorsal dominance (Voorn et al., 2004). Progression from medial to lateral (VTA to SNc) reflects a shift from prediction error encoding to salience encoding (Bromberg-Martin et al., 2010). **B.** Recording locations and connectivity of NAc, DMS and DLS. The coronal section shown is taken from Paxinos and Watson (2007) and are approximations from recording sites from studies shown in Chapters 2-4. Recording sites for nucleus accumbens core (NAc) were ~1.6 mm anterior to Bregma, 1.5 mm lateral to midline, and 6 mm ventral to brain surface. Recording sites for dorsal lateral striatum (DLS) were ~1.0 mm anterior to Bregma, 3.2-3.6 mm lateral to midline, and 3.5 mm ventral to brain surface. Abbreviations: AI, agranular insula; OFC, orbitofrontal cortex; PFC, prefrontal cortex; ACC, anterior cingulate cortex; SMC, sensorimotor cortex; NA, nucleus accumbens; DMS, dorsal medial striatum; DLS, dorsal lateral striatum; VTA, ventral tegmental area; SNc, substantia nigra pars compacta; DA, dopamine.

reward than expected—positive prediction error; less reward than expected—negative prediction error; reward fully expected—no prediction error) (Schultz, 2006; Bromberg-Martin et al., 2010). It is also known that DA response to rewards shift to cues that predict rewards (Schultz, 2006, 2013). These teaching signals indicate value-related responses that inform behavioral policies to obtain the best outcomes and avoid negative outcomes (Schultz, 2006; Haber and Knutson, 2009; Bromberg-Martin et al., 2010). This type of coding in DA neurons has been referred to as motivational value coding, where rewarding events cause excitation and aversive events cause inhibition of DA neurons (Bromberg-Martin et al., 2010).

In addition to providing reinforcing teaching signals and encoding motivational value, recent evidence shows that phasic-bursting DA neurons also encode motivational salience. DA neurons encoding motivational salience are excited by both rewarding *and* aversive events. This is unlike motivational value coding, where this type of DA neuron is only excited by rewarding events and inhibited by aversive events (Bromberg-Martin et al., 2010). A recent study found evidence of midbrain DA neurons encoding motivational value *and* motivational salience in monkeys performing a Pavlovian task where rewards and aversive outcomes were presented (Matsumoto and Hikosaka, 2009). Specifically, Matsumoto and Hikosaka (2009) found that dopaminergic neurons were excited by both rewarding events (liquid reward) *and* aversive events (air-puffs), and that this neural response pattern was also true for stimuli predicting these events. Furthermore, they found evidence that anatomical divisions of DA neurons in the midbrain were roughly arranged by

what they encode and what their targeted projections are, which is explained further below.

Major cortical and sub-cortical sources of dopaminergic innervation originate in the midbrain in the ventral tegmental area (VTA) and the substantia nigra pars compacta (SNc) (Björklund and Dunnett, 2007; Matsumoto and Hikosaka, 2009; Bromberg-Martin et al., 2010; Parker et al., 2016). DA neurons encoding motivational value and motivational salience are anatomically distributed across a topographical gradient in the VTA and SNc, with motivational value signals found in DA neurons in the ventromedial SNc and lateral parts of the VTA and motivational salience signals found in DA neurons in the dorsolateral SNc (Matsumoto and Hikosaka, 2009; Bromberg-Martin et al., 2010).

As mentioned above, DA neurons provide projections to many important targets in the reward circuit. Of particular relevance are the major projections to the striatum. It has been hypothesized that the topographically organized gradient of DA neurons in the midbrain also corresponds to striatal projections and the functional organization of the striatum (Bromberg-Martin et al., 2010; Parker et al., 2016). DA neurons located more ventromedially in the SNc and VTA that signal motivational value project to ventral regions of striatum, and DA neurons located more dorsolaterally in the SNc that signal motivational salience project more prominently to dorsal regions of striatum (Matsumoto and Hikosaka, 2009; Bromberg-Martin et al., 2010).

Indeed, Parker and colleagues (2016) found evidence that DA neurons that project to specific regions of striatum encode specialized information that would

correspond to this topographical organization of projections from the midbrain. Specifically, they recorded from striatal terminal regions of midbrain DA projections from the VTA/SNc to DMS and from the VTA/SNc to NAc in mice performing a probabilistic reversal lever-press task. They found that VTA/SNc to DMS projection neurons preferentially fired when the animal was about to make a contralateral choice and that VTA/SNc to NAc projection neurons preferentially fired to reward (Parker et al., 2016). This fits with the DMS being involved in evaluation of and generating actions and the NAc being involved in processing of reward (Graybiel et al., 1994; Ikemoto and Panksepp, 1999; Roitman et al., 2005; Balleine et al., 2007; Tai et al., 2012; Parker et al., 2016).

Exposure to drugs of abuse has long been known to affect these brain regions and the dopaminergic system that make up reward circuitry and translate to subsequent maladaptive actions and behaviors commonly seen in addiction (Everitt and Robbins, 2005; Robbins et al., 2008; Stalnaker et al., 2009; Koob and Volkow, 2010). Research has shown that addicts have a difficult time controlling impulses and decision-making, which can lead to poor choices and impaired judgment when it comes to seeking and taking more drugs and allowing their addiction to spiral out of control (Volkow and Baler, 2015; Volkow and Morales, 2015). Understanding neural correlates underlying normal and maladaptive behaviors in learning and decision-making is imperative in moving forward with treating substance abuse effectively.

Previous research has shown that reward-based learning and decision-making can be under the control of goal-directed and habit-like processing within reward circuitry regions discussed above (Balleine and O'Doherty, 2010). Specifically, goal-

directed behaviors are under the control of value-related response-outcome associations (R-O; e.g., "If I do this, I will get a reward"), which fall under the control of ventral regions such as the NAc and more dorsal regions of striatum such as the DMS (Balleine and O'Doherty, 2010; Burton et al., 2014a, 2015). On the other side of the spectrum, behaviors related to habitual processing and responding after extended periods of learning are under the control of stimulus-response associations (S-R; e.g., "When I see this cue, I automatically respond"). Neural correlates related to these behaviors are found in more lateral regions of dorsal striatum (DLS) (Balleine and O'Doherty, 2010; Burton et al., 2015). Following is a more detailed discussion of behavioral and neural correlates amongst these sub-regions of striatum.

A more detailed look within the reward circuit: the striatum

The striatum is a critical input component to the basal ganglia, which is a complex set of subcortical brain structures involved in the selection and execution of actions (Graybiel et al., 1994; Kreitzer and Malenka, 2008; Matamales et al., 2009). More specifically, the striatum is known to collect and process emotional, cognitive, and sensorimotor information from the cortex and midbrain to generate or suppress actions (Graybiel et al., 1994; Matamales et al., 2009). More than 95% of all neurons in the striatum are inhibitory (meaning they release the neurotransmitter gamma-aminobutyric acid, or GABA) medium spiny neurons (MSNs), with the remaining neurons being other types of inhibitory interneurons (Matamales et al., 2009). Striatal MSNs receive excitatory glutamatergic synapses from multiple cortical and thalamic inputs and dopaminergic innervation from the midbrain. There are two types of

MSNs that are equally distributed and intermingled throughout the striatum: those that express dopaminergic D1-type receptors (D1Rs) and those that express dopaminergic D2-type receptors (D2Rs) (Graybiel et al., 1994; Kreitzer and Malenka, 2008; Matamales et al., 2009).

For the most part, neurons expressing D1Rs and D2Rs project to the direct and indirect motor pathways of the basal ganglia, respectively. The D1R direct pathway facilitates body movements, while the D2R indirect pathway suppresses body movements (Albin et al., 1989; Gerfen et al., 1990; Bromberg-Martin et al., 2010; Hikida et al., 2010). It has been theorized that tonic bursts of DA that produce high DA concentrations (positive reward prediction errors) activate D1Rs that project to the direct pathway to produce motor movements and obtain high value outcomes (Frank, 2005; Hikosaka, 2007; Bromberg-Martin et al., 2010). On the other hand, pauses in tonic bursting of DA neurons (negative reward prediction errors) that produce low DA concentrations inhibit D2Rs that project to the indirect pathway to suppress motor movements and avoid low value outcomes (Frank, 2005; Hikosaka, 2007; Bromberg-Martin et al., 2010). Similar divisions of labor are evident both in the ventral and dorsal striatum (Grace et al., 2007; Bromberg-Martin et al., 2010; Lobo et al., 2010). However, it is important to note that there is recent evidence suggesting that there is not such a clear division of D1- and D2-expressing neurons in the NAc projecting to the direct and indirect output nuclei of the basal ganglia as once originally thought (Kupchik et al., 2015).

Overall, evidence suggests that the majority of the cell-type in the striatum is GABAergic MSNs, and that they receive both glutamatergic and dopaminergic

innervation. MSNs can express two major DA receptor sub-types that affect subsequent motor output. Below is a brief discussion of cellular mechanisms, synaptic plasticity, and how addiction affects cell-types in these regions, followed by a description of what kinds of relevant information single neurons are encoding in the striatum.

Synaptic plasticity in the striatum after cocaine exposure

Cocaine acts upon the dopamine system by blocking the dopamine transporter (DAT), located on the pre-synaptic membrane of DA neurons projecting to MSNs of the striatum. DAT blockage prevents reuptake of DA from the synapse, flooding it with an overabundance of DA and subsequently affecting both pre- and post-synaptic mechanisms. Seminal work done by Malenka, Bonci, and colleagues (Ungless et al., 2001) showed synaptic changes in MSNs in the striatum after just one exposure to cocaine via injection in animals. They were able to pharmacologically isolate MSN glutamatergic receptor subtypes (AMPAR and NMDAR) and their components to excitatory post-synaptic currents in MSNs of the striatum. They found robust long-term potentiation (LTP) after a single injection, with a particular enhancement in the AMPAR contribution to the EPSC (Ungless et al., 2001).

They also found that this single exposure strengthening of the synapse also occludes further LTP, saline injected animals reached higher EPSC currents when high frequency stimulation was induced, whereas cocaine animals stayed at a steady EPSC current even after high frequency stimulation was induced (Ungless et al., 2001). The AMPAR/NMDAR ratio is enhanced for up to five days after a single cocaine exposure. With only one exposure to the drug, we can see that this short-term

drug exposure is powerful enough to begin the strengthening of synapses and connections in this area.

Research shows that strengthening of synapses occurs even after very short-term use of cocaine, and that there must be long-term changes that follow long-term usage as well. After repeated cocaine exposure, there is a significant increase in the number of spines (spinogenesis) in cocaine animals compared to controls (Nestler, 2013). However, some of these spines may be "silent," in that the presence of functioning NMDARs can be detected, but there is a lack of AMPARs. This means that no depolarization of the neuron can occur to remove the magnesium block on NMDARs, and does not allow for activation of these receptors at resting state (Busetto et al., 2008; Wolf and Ferrario, 2010). These silent spines may incorporate functioning AMPARs later on after cessation of the drug and during withdrawal, and indeed, work done by Marina Wolf showed that both MSNs in the striatum and neurons in the VTA do scale up AMPARs after termination of drug use (Conrad et al., 2008).

Wolf's data also showed that there is a subunit switch in the AMPARs that is inserted during withdrawal from drugs, specifically lacking the GluR2 subunit making these AMPARs (GluR1/3) even more permeable to calcium (Conrad et al., 2008; Wolf and Ferrario, 2010). This is the phenomenon of AMPAR scaling of synapses that are "hungry" for input following termination of drug use. In addition, Lee and colleagues showed where silent synapses were generated in the NAc after an extended period of cocaine exposure, which also showed a similar pattern in maturing

during withdrawal to incorporate the more calcium permeable GluR2-lacking AMPARs (Lee et al., 2013).

Another change in synaptic mechanisms after long-term drug use is in the epigenetic realm of protein modulation of addiction (Robison and Nestler, 2011). Specifically, the delta fos B (ΔFosB) protein encoded by the FosB gene, which is an immediate early gene that is rapidly transcribed in response to particularly salient stimulation, is enhanced in neurons that have been exposed to addictive drugs of abuse. It has been reported that all Fos family proteins are induced by drug exposure, and particularly ΔFosB after chronic administration of drugs (Robison and Nestler, 2011). In brief, this mechanism starts with repeated drug use signaling through G-protein coupled DA receptors that then activate other downstream substrates (protein kinase A, or PKA) to stimulate transcription, which induces production of ΔFosB, which has a very long half-life because of the absence of a degradation protein (PEST). This induction has been reported to only occur in NAc MSNs that express D1 dopamine receptors and it can persist for weeks after terminating drug use. Nestler also produced seminal work that showed that ΔFosB overexpression in NAc MSNs increased many behavioral responses to cocaine (Kelz et al., 1999; Colby et al., 2003; Zachariou et al., 2006).

Overall, synaptic enhancements or strengthening of connections that are present early after drug use seem to persist through long-term use of drugs of abuse. The epigenetic changes seen in long-term drug abuse may mediate the final stages of addiction through the transcription of Fos genes. Drugs of abuse are known to affect reward responses in DA neurons and this may reflect an aberrant learning mechanism

that is mediated in the later stages of addiction by reducing DA levels that may motivate users to seek more drugs or valued outcomes to enhance DA release. These learning mechanisms and responses to valued outcomes are discussed in further detail below.

Encoding in single neurons within striatum

Although significant progress has been made in understanding how drugs of abuse, specifically cocaine, impact cellular mechanisms and synaptic plasticity, we are still trying to understand how encoding in single neurons during decision-making changes after drug exposure. Through many single-neuron and interference studies it has become clear that the striatum is critical for reward-guided and habitual behaviors (Burton et al., 2015; Bissonette and Roesch, 2016).

The striatum has been previously termed as an "actor-critic" in models of reinforcement learning, with ventral regions of striatum acting as the "critic" in providing value information in the form of prediction errors necessary for updating subsequent actions controlled by dorsal regions of striatum, or the "actor" in the model (Barto, 1995; Sutton and Barto, 1998; Joel et al., 2002; Redish, 2004a; Niv and Schoenbaum, 2008; Takahashi et al., 2009; Padoa-Schioppa, 2011; Van der Meer and Redish, 2011; Bissonette and Roesch, 2016). There are many overlapping neural correlates in the striatum related to reward-guided decision-making that fire to cues predicting reward, in anticipation of reward, during reward delivery, and during action initiation to obtain rewards (Burton et al., 2015; Bissonette and Roesch, 2016). Overall, there seems to be a trend with more reward and value-related information being processed in ventral striatum and more associative and motor-related

information being processed moving dorsally and laterally through striatum (Burton et al., 2015; Bissonette and Roesch, 2016). The following sections will discuss in more detail what single neurons are encoding in these regions related to reward-guided decision-making.

The ventral striatum

In addition to being characterized as the "critic" in the actor-critic model of reinforcement learning, the ventral striatum (VS) has also been termed the "limbic-motor interface," as it processes multiple inputs from cortical and midbrain regions to motivate behaviors to both appetitive and aversive stimuli in the environment (Bissonette and Roesch, 2016). VS, and in particular the nucleus accumbens core (NAc) region of VS, receives many glutamatergic inputs from cortical areas like the OFC, which is known to signal predictions about future outcomes. Combined with dopaminergic innervation in the form of reward prediction errors, the NAc processes this information about upcoming and previous outcomes to generate motivated behaviors to approach and avoid important stimuli in the environment (Bromberg-Martin et al., 2010; Bissonette and Roesch, 2016). These inputs and subsequent neural correlates likely contain important value-related or motivation-related information that needs to be projected to motor areas that are critical for action selection.

There is ample evidence to support these theories from experiments implementing pharmacological manipulations and single-unit recordings of NAc in a variety of tasks performed by both rats and monkeys showing that activity in NAc is modulated by cues that predict reward after an instrumental response is performed (as

reviewed in Bissonette and Roesch, 2016). These include reward-seeking, cost-benefit analysis, and delay/effort discounting tasks and a variety of instrumental tasks including go/no-go, lever pressing, discrimination, maze running, and eye movement paradigms (Burton et al., 2015; Bissonette and Roesch, 2016). Again, the overall general finding from these studies was activity in NAc was modulated by reward-predicting cues, and this activity may reflect value or motivation.

To further investigate this question, we implemented a novel paradigm where rats were motivated by both an appetitive reward (sucrose solution) and an aversive threat of punishment (quinine solution) that were signaled by odor cues (Bissonette et al., 2013). Following presentation of odor cues signaling the outcome of that particular trial, rats were then presented with an instructional light cue directing them to adjacent fluid wells to obtain rewards and avoid threats of punishment upon correct completion of the trial. Importantly, in this study the odor cues were presented *before* the instructional light stimulus to dissociate value signals from specific motor planning (Burton et al., 2015; Bissonette and Roesch, 2016).

We found that rats were highly motivated when cues predicted a possible large reward or a possible quinine punishment, as evidenced by faster reactions times and increased accuracy on these trial types compared to neutral cues predicting a small sucrose reward (Bissonette et al., 2013; Burton et al., 2015; Bissonette and Roesch, 2016) . Furthermore, we found NAc neurons that fired strongly for cues that predicted large rewards and weakly for quinine risk odor cues compared to neutral cues; these neurons represented value. Other NAc neurons fired strongly to both cues that predicted large reward and threatened punishment; these neurons signaled

motivation. Overall, there were two populations of neurons that encoded value and motivation separately which fits with the NAc as being both the evaluative "critic" in the actor-critic model and a "limbic-motor interface" by integrating value and motivational information (Bissonette et al., 2013; Burton et al., 2015; Bissonette and Roesch, 2016.

In the odor-guided choice task used in the experiments in this dissertation (explained in detail in the next Chapter section), we have previously shown that the firing of NAc neurons was modulated by the value of a reward associated with a particular direction (Roesch et al., 2009; Burton et al., 2015). This task differs from the previous task in that the odor cues predicted differently valued rewards *and* the direction necessary to respond to obtain the reward. The majority of neurons fired significantly more strongly for cues that predicted high-value outcomes for actions made in a particular direction, or into the cell's preferred direction or response field (Roesch et al., 2009; Burton et al., 2015). This activity was also correlated with faster reaction times, which fits with neurons in the NAc signaling motivational value (Roesch et al., 2009; Burton et al., 2015). In each of these studies, we concluded that the NAc encodes predicted value of cues, either in conjunction with or in the absence of response information. This is in line with its roles as both the critic in the actor-critic model of reinforcement learning and as the limbic-motor interface.

Importantly, single neurons in the NAc also encode the value of expected outcomes by increasing activity *after* instrumental responses or Pavlovian cues as animals anticipate reward delivery (Bowman et al., 1996; Hassani et al., 2001; Cromwell and Schultz, 2003; Setlow et al., 2003; Janak et al., 2004; Nicola, 2007;

Khamassi et al., 2008; Van der Meer and Redish, 2009; Ito and Doya, 2009; Roesch et al., 2009; Cai et al., 2011; Tanaka et al., 2016). This outcome expectancy firing in NAc plays a critical role in timing rewarded outcomes during performance of behavioral tasks, including temporal discounting tasks (Meck et al., 2008; Singh et al., 2011; Mello et al., 2015; Takahashi et al., 2016). Further, it has been shown that loss of NAc is detrimental to behavior on temporal discounting tasks (Cousins et al., 1996; Cardinal et al., 2001, 2004; Bezzina et al., 2007; Floresco et al., 2008; Kalenscher and Pennartz, 2008; McDannald et al., 2011; Burton et al., 2014a) and alters neural signaling in downstream regions such as the VTA and dorsal striatum (DS) during task performance (Burton et al., 2014a; Takahashi et al., 2016).

Taken together, it is clear from these studies that the NAc is critical in signaling value, motivation, and reward expectancy to integrate and provide information to downstream areas to guide decision-making. Research suggests that the NAc is one of the first regions targeted by drugs of abuse because of its strong dopaminergic innervation from the VTA and its significant role in positive reinforcement (Koob and Volkow, 2010; Steinberg et al., 2014; Keiflin and Janak, 2015; Lüscher, 2016). The NAc sends projections back to the VTA and other midbrain regions, which then project to more dorsal regions of striatum forming a spiraling circuitry that is important for optimal action selection, which is known to be impacted by drugs of abuse and critical for drug seeking (Haber et al., 2000; Belin and Everitt, 2008; Haber and Knutson, 2009; Everitt and Robbins, 2013; Burton et al., 2015). Next, we consider what single neurons are encoding in dorsal striatum.

The dorsal striatum

We know that single neurons in the NAc signal value, motivation, and anticipatory signals related to future outcomes, as discussed above. Furthermore, we know that NAc communicates with the DS indirectly via DA neurons in the midbrain. The DS, in particular DLS, is thought to be the "actor" in the actor-critic model. We and others have found highly associative motor-related signals in DS, but signals are not simply integrated and transformed into motor output. There are many reward-related and stimulus-driven behaviors and overlapping neural correlates, with a trend of signaling appearing along the diagonal of DS (Bissonette and Roesch, 2016). This appears to be a basic trend from ventral-medial to dorsal-lateral where more reward-related neural correlates are found in ventral-medial regions and more associative and motor-related neural correlates are found in dorsal-lateral regions. This has been demonstrated in single-neuron recording studies in the primate caudate, which is most similar to DMS in rodents (Nakamura et al., 2012; Burton et al., 2015; Bissonette and Roesch, 2016).

Nakamura and colleagues (2012), used an eye movement paradigm where monkeys had to respond directionally left or right after fixation in a blocked-trial design where one direction produced a large reward and the other produced a small reward. They found that single neurons within more ventral-medial regions of caudate (DS) were more strongly driven by larger reward independent of the actions necessary to obtain it, whereas single neurons in more dorsal-lateral regions of caudate exhibited high directionality and reward selectivity, but had no preference for

large over small reward (Nakamura et al., 2012; Burton et al., 2015; Bissonette and Roesch, 2016).

Overall, this fits with the trend mentioned above with more ventral-medial regions having more reward-related correlates and more dorsal-lateral regions being more associative; other studies in monkeys and rats have supported topology(Apicella et al., 1991; Hollerman et al., 1998; Takahashi et al., 2007; van der Meer et al., 2010; Cai et al., 2011). In monkeys, work by Schultz and colleagues examining task-related activity throughout monkey striatum, including dorsal and ventromedial parts of putamen (DLS), dorsal and ventral caudate (DMS), and in NAc as monkeys performed a go/no-go task for reward (Apicella et al., 1991; Hollerman et al., 1998) has shown that twice as many neurons that increased firing to reward for both go and no-go trials in ventral compared to dorsal striatal areas, whereas the number of neurons that were modulated by movement type, including non-rewarded movements (i.e., no-gos), were more prominent in the head of the caudate (Apicella et al., 1991; Hollerman et al., 1998).

To summarize, along a diagonal gradient within DS we see more reward-related responses in ventral-medial regions and more motor-related responses in dorsal-lateral regions of DS. Across striatum there appears to be an integration of overlapping information from ventral regions and dorsal regions in generating actions to obtain valued outcomes. Next, we consider what these correlates represent during learning and behavior from the perspective of representing goals and habits in medial and lateral aspects of DS.

Reward-related actions are thought to reflect the integration of two different learning processes, one that controls goal-directed behavior and the other the acquisition of habits (Balleine and O'Doherty, 2010). Goal-directed behavior is governed by neural representations of associations between responses and outcomes (R-O), whereas habit formation is controlled by relationships between stimuli and responses (S-R). One way to assess the role that specific brain areas play in these two functions is to devalue the reward that the animal is anticipating. If behavior is under control of the expected outcome (R-O), then that behavior should cease when the outcome is no longer desirable after devaluation. However, if behavior is stimulus-driven (S-R), then outcome devaluation should not impact behavioral output. This procedure has demonstrated that both NAc and DMS lesions impair goal-directed behavior, leaving the animal's behavior stimulus-driven and habitual. In contrast, DLS lesions inhibit the formation of habits, leaving the animal's behavior under the control of anticipated goals (Balleine and O'Doherty, 2010).

Consistent with these findings, Gremel and Costa (2013) have recently shown that neural activity in both DMS and DLS carry signals related to both goal-directed behavior and habitual action, respectively. Their task allowed for a clean dissociation between goal- and stimulus-driven behavior by training animals on either a random ratio or random-interval schedule of reinforcement, respectively. Reinforcer devaluation confirmed that behaviors under these schedules were indeed goal- and habit-driven (Gremel and Costa, 2013). Consistent with previous work, lesions to DMS hindered goal-directed decisions promoting habitual behavior during performance of their task, whereas lesions to DLS hindered habitual responding,

forcing animals to be goal-driven. In both cases, animals could still perform the task, but were using different functional mechanisms to control behavior. These results suggest that neural correlates that control goal-directed and habitual actions are processed in parallel, and that these regions are modulated differently depending on the mechanism that is ultimately controlling behavior.

All of these studies provide clear evidence that there are parallel and differential firing patterns throughout regions of DS—along the diagonal from ventral-medial to dorsal-lateral DS and between DMS and DLS. These overlapping correlates suggest a complex integration of information as neural activity propagates through the striatum. Furthermore, these studies demonstrate that the DS carries signals related to goal-directed and habitual responding, but the strength of modulation on these task correlates depends on the mechanism currently governing behavior (goal or habit).

In summary, consistent with the "actor-critic" model of reinforcement learning, the sub-regions of striatum evaluate upcoming outcomes in order to motivate and generate behaviors to obtain the best outcomes. As discussed above, many studies show NAc single-neurons and populations of neurons that signal value, motivation, and reward expectancy. As we move upstream in the striatum to ventral-medial regions and DMS, we see more associative encoding related goal-directed behaviors, reward, and the responses made to obtain them. Finally, as we move up- and out-ward to DLS, many studies show strong associative encoding related to motor-responses in various tasks. The communication amongst these regions through

indirect connections with midbrain DA neurons allows the proliferation of information throughout the striatum.

Although the NAc, DMS, and DLS are not directly connected, more ventral regions likely impact more dorsal regions on a diagonal trend through the spiraling circuitry with DA neurons (as discussed above in previous section) to allow the propagation of information from more limbic regions to associative and sensorimotor networks to control behavior (Burton et al., 2015; Bissonette and Roesch, 2016). This allows for behaviors to become more stimulus-driven and efficient, which is consistent with the idea that at first behaviors are goal-directed and over time with extended learning and training become more habitual in nature (Bissonette and Roesch, 2016). This would also fit with current theories of addiction as loss of goal-directed behavioral control under control of more ventral regions of striatum in favor of more habit-like processing and responding to cues in the environment that are under control of more dorsal regions of striatum (Everitt and Robbins, 2005, 2013, 2016; Hyman et al., 2006; Belin and Everitt, 2008).

Throughout this section, I discussed neural correlates in normally functioning circuits irrespective of drug use. It is critical to understand how these circuits process and output information in normal behavior before we can begin to understand how the structure and function of the circuit changes after exposure to drugs of abuse. In understanding the normal function of the striatum and the subsequent changes when the circuit is perturbed, we can better understand and improve treatment options in humans suffering this loss of behavioral control in this disorder. The theory of

addiction as a loss of goal-directed control in favor of habitual control and how we can specifically study this in animal models is discussed in the following section.

Addiction, decision-making, and how to investigate it in animal models

Addiction afflicts millions of people and there have been many human imaging studies that have used functional magnetic resonance imaging (fMRI) to try to elucidate the neural mechanisms underlying some of the behavioral and neural changes that occur in this disorder in humans. In particular, these studies have looked at response patterns related to drug-related stimuli and to systemic administration of drugs of abuse while in the scanner that have shown the negative impact of addiction in many regions of the brain (Volkow et al., 2004, 2015; Delgado, 2007; Fowler et al., 2007). Specifically, regions involved in processing reward, drug cues and stimuli and feelings of craving such as frontal cortex, striatum and ventral tegmental area (where dopamine projections originate) have been implicated. The negative impact that drugs of abuse have on these regions seems to correlate well with common cognitive deficits in addicts such as: impairments in response inhibition and cognitive flexibility in reinforcement learning tasks and reversal learning, decision-making and impulsiveness in gambling tasks and delay discounting tasks, and working memory deficits (Robbins et al., 2008; Volkow and Baler, 2015; Everitt and Robbins, 2016).

Even though we have learned much from human studies, there are important limitations to consider when trying to understand changes in neural activity underlying maladaptive behaviors seen in addiction. One major limitation when considering results from the human addiction literature are individual differences,

which can arise from genetic predispositions, environmental or socioeconomic factors, different stages of addiction, or any combination of these. Importantly, the relevance between reward circuitry in humans and other animals, in particular primates and rodents, has been established throughout the literature (Haber and Knutson, 2009; Balleine and O'Doherty, 2010). Notably, the focus of the proposed research, the striatum as a whole and the sub-regions of striatum in humans have primate and rodent homologs that display similar activities in response to reward-based learning and decision-making (O'Doherty, 2004; Haber and Knutson, 2009; Balleine and O'Doherty, 2010; Liljeholm and O'Doherty, 2012).

The ventral regions of striatum in human anatomy correspond well to ventral regions of rodent striatum, and they are both implicated in motivation (Balleine and O'Doherty, 2010; Liljeholm and O'Doherty, 2012). Increases in VS activity have been attributed to motivation, through review of many fMRI studies (Liljeholm and O'Doherty, 2012). VS activity is also seen in imaging studies to increase in anticipation of reward and be scaled by reward magnitude (O'Doherty, 2004; Haber and Knutson, 2009). The DS in humans and primates consists of the caudate nucleus and the putamen which corresponds to DMS and DLS in rodents, respectively (Haber and Knutson, 2009; Balleine and O'Doherty, 2010). Important human imaging studies have attributed activity in DMS/caudate to goal-directed behaviors and activity in DLS/putamen to habitual behaviors (Delgado, 2007; Liljeholm and O'Doherty, 2012). Knowing that we see similar reward-related activity in similar brain regions between humans, primates, and rodents is significant in extending

research to useful animal models of addiction to look at changes in brain and behavior.

In addition to showing similar circuitry and response patterns, there are also similar behavioral task versions used in animal research to understand common cognitive deficits seen in tasks used human studies. Specifically, addicts show deficits and impulsive behavior in delay discounting tasks (Heil et al., 2006; Reynolds, 2006; Mackillop, 2013; Volkow and Baler, 2015), which measure the value of reward in regards to its delayed receipt (Odum, 2011). When given a choice, most human beings would choose to have an immediate reward over a delayed reward. However, increasing the value of the delayed reward may have a significant impact on future choices. Delay discounting gives the choice of an immediate small reward versus a larger delayed reward, referred to as the impulsive choice versus the self-controlled choice, respectively (Odum, 2011). It has been shown that cocaine addicts have a difficult time delaying future rewards and often impulsively choose the more immediate reward more frequently even when it is not in their best interests (Heil et al., 2006).

Many advances have been made in research involving animal models of addiction as well. In the past, a common way to model addiction in animals was to systemically inject drugs of abuse into the animal and then measure subsequent behavior and/or neural activity. Now, instead of injections, intravenous self-administration models have been developed in order to allow the animal to seek and take drugs of abuse upon their own volition (Thomsen and Caine, 2005). Many researchers feel that this is a better model that correlates better to human conditions,

as the animal has the choice of how much and how often to consume the drug, and they are in direct control of the drug entering their system. These models commonly feature rats lever-pressing or nose-poking to deliver drugs via catheter directly into the circulatory system.

In the literature, addiction theories have emerged that indicate changes in behavior are due to habitual behaviors taking over goal-directed decision-making. Specifically, the locus of behavioral and neural control moves from more ventral regions of striatum to more dorsal regions of striatum as addiction progresses (Everitt and Robbins, 2005, 2013, 2016; Hyman et al., 2006; Belin and Everitt, 2008). It has been hypothesized that the NAc and the downstream neural projections between VS and DS are one of the first to be targeted by drugs of abuse (Belin and Everitt, 2008; Everitt and Robbins, 2013). Anatomical studies have shown that these spiraling dopamine projections from VS to midbrain, back to striatum, cascading upwards to DS form an important part of the reward circuit discussed earlier (Haber et al., 2000; Belin and Everitt, 2008; Haber and Knutson, 2009; Burton et al., 2015).

Even though many advances have been made in studying the impact of drugs of abuse on behavior and brain, it is still unclear how single neurons, whole brain regions, communication amongst these regions, and subsequent decision-making on complex behavioral tasks are affected after drug exposure. The focus of my research is on striatal areas within reward circuitry that are hypothesized to be among the first areas affected by drug exposure (Keiflin and Janak, 2015; Lüscher, 2016) (Figure 1B). The objective of my research was to determine the long-term impact of previous cocaine self-administration on striatal signaling during performance of a complex

behavioral task that often-changed expected value contingencies. My research used an animal model of exposure to drugs of abuse, specifically intravenous cocaine self-administration in rats, to examine the impact on behavior and neural signaling in striatal areas during performance of a reward-based decision-making task. I used a well-established modified version of a delay discounting task described in the next section of this chapter.

We recorded from single neurons in NAc and DLS in motivated rats that had prior cocaine self-administration experience while performing an odor-guided decision-making task to obtain liquid sucrose rewards. We found evidence that cocaine biased behavior towards higher-valued outcomes and that neural encoding in both NAc and DLS was altered both at the single-neuron and neuron population level within each region. After cocaine exposure, we saw differences in counts of neurons that increased and decreased activity during different task epochs in both NAc and DLS. This led us to ask what would happen to signaling if we disrupt the circuit between NAc and DLS. The last goal of my research was to determine the impact of pharmacological lesions of the NAc on signaling in DLS while rats perform the same decision-making task. When combined, these data will help us understand what nodes in this circuit are disrupted after chronic drug use, and the impact this disruption has on downstream brain areas and decision-making with implications for treatment strategies.

Odor-guided decision-making task: a modified version of a delay-discounting task

The basic idea of standard delay discounting tasks is to look at a subject's or animal's preference in choice between an immediate small reward and a delayed large reward (Reynolds, 2006; Odum, 2011; Roesch and Bryden, 2011). It is known that humans and animals prefer an immediate reward over a delayed reward, even when the delayed reward is more economically valuable in the long run (Herrnstein, 1961; Ainslie, 1974; Thaler, 1981; Kahneman and Tversky, 1984; Rodriguez and Logue, 1988; Loewenstein and Prelec, 1992; Richards et al., 1997; Evenden and Ryan, 1999; Ho et al., 1999; Cardinal et al., 2001; Mobini et al., 2002; Winstanley, 2004; Kalenscher et al., 2005; Kalenscher and Pennartz, 2008; Ballard and Knutson, 2009). As the delays to the larger reward become longer we know that animals often abandon delayed rewards and bias choices toward the more immediate smaller reward. This change in choice preference can be looked at as a measure of impulsivity, since over the course of many trials it would be more economically beneficial to wait for the larger reward (Roesch and Bryden, 2011). It is important to note that these standard delay-discounting tasks manipulate both the delay to and size of reward during the same trial.

Many studies have looked at pharmacological and anatomical manipulations and subsequent performance on delay discounting tasks, but few have looked at neural correlates underlying decision-making and behavior in these types of tasks. I have implemented a modified version of a delay-discounting task where we independently manipulate the delay to and size of reward in different and separate blocks of trials (Roesch et al., 2006, 2007a, 2009; Calu et al., 2010; Stalnaker et al.,

2010; Bryden et al., 2011a). In doing this, our goal has been to determine how delay and magnitude can impact neural encoding independently from each other in different brain areas known to be involved in reinforcement learning and decision-making and how this may be impacted by certain diseases and disorders (Roesch and Bryden, 2011). A detailed description of the task is below.

Experimental subjects are male Long Evans rats obtained from Charles River Laboratories. All training and recording sessions during performance of this odor-guided task were conducted in aluminum chambers approximately 18" on each side with downward sloping walls narrowing to an area of 12" x 12" at the bottom. On one wall, a central nose port was located above two adjacent fluid wells. Two houselights were located above this panel. Odors were delivered directly into the nose port via a hemicylinder and magnetic solenoid valve located behind the port. Liquid sucrose was delivered via magnetic solenoid valves to fluid wells upon correct responses only. Houselights were illuminated at the start of the trial and were extinguished upon the animal leaving the fluid well at the end of the trial. Task control was implemented via computer programming (C++). Odor port and fluid well entries and exits were monitored by disruption of photo beams.

The basic design of a trial is illustrated in Figure 2A. Each trial began by illumination of the houselights that informed the animal it could make a nose-poke into the central nose port. Once the rat broke the photo beam entering the nose port, it had to wait for a 500ms period before the odor cue was delivered for another 500ms. One of three different odors (2-Octanol, Pentyl Acetate, or Carvone) was delivered to the port on each trial. One odor instructed the rat to go to the left fluid well to receive

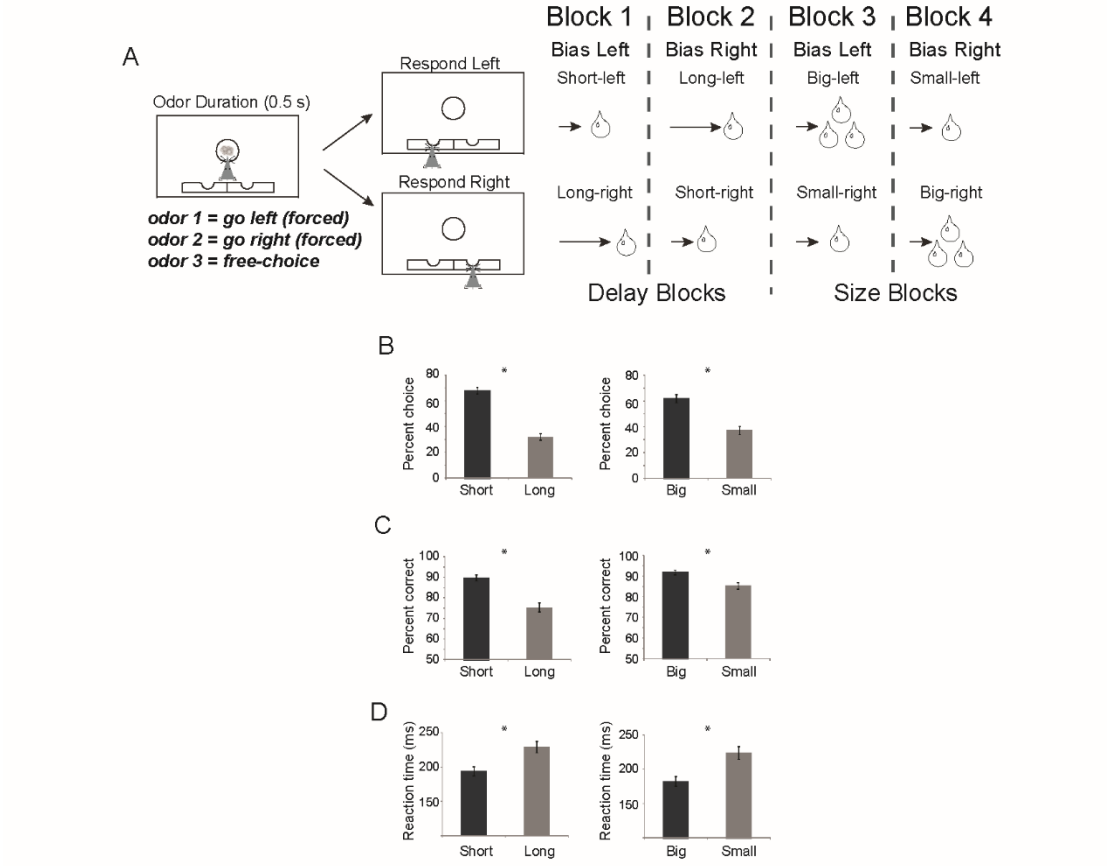

Figure 2: Task and behavior **A.** The sequence of events in each trial block. For each recording session, one fluid well was arbitrarily designated as short (a short 500ms delay before reward) and the other designated as long (a relatively long 1-7s delay before reward) (Block 1). After the first block of trials (~60 trials), contingencies unexpectedly reversed (Block 2). With the transition to Block 3, the delays to reward were held constant across wells (500ms), but the size of the reward was manipulated. The well designated as long during the previous block now offered 2-3 fluid boli whereas the opposite well offered one bolus. The reward stipulations again reversed in Block 4. **B**. The impact of delay length (left) and reward size (right) manipulations on choice behavior during free-choice trials. **C**. Impact of value on forced-choice trials percent correct for delay (left) and size (right) manipulations. **D.** Reaction times (odor offset to nose unpoke from odor port) on forced-choice trials (expressed in ms) comparing short vs. long delay trials and big vs. small reward trials. (t test, $p<0.05$). Error bars indicate SEMs.

reward (forced-choice), a second odor instructed the rat to go to the right fluid well to receive reward (forced-choice), and a third odor indicated that the rat could obtain reward at either well (free-choice). Odors were counterbalanced across rats. The meaning of each odor did not change across sessions. Odors were presented in a pseudorandom sequence such that the free-choice odor was presented on 7/20 trials and the left/right odors were presented in equal proportions.

During training and recording, one well was randomly designated as short (500ms delay to reward) and the other long (1-7s delay to reward) at the start of the session. In the second block of trials, these contingencies were switched. The length of the delay under long conditions abided by the following algorithm: the side designated as long started off as 1s and increased by 1s every time that side was chosen on a free-choice odor (up to a maximum of 7s). If the rat chose the side designated as long less than 8 out of the previous 10 free choice trials, the delay was reduced by 1s for each trial to a minimum of 3s. The delay to reward for long forced-choice trials was yoked to the delay in free-choice trials during these blocks. In later blocks, the delay preceding reward delivery was held constant (500ms) while manipulating the size of the expected reward. The reward was a 0.05 ml bolus of 10% sucrose solution. For big reward, an additional bolus was delivered 500ms after the first bolus. Rats performed at least 60 correct trials per block before the block was switched, for a total of approximately 240 trials per session. Essentially there were four basic trial types (short, long, big, and small) by two directions (left and right) by two stimulus types (free- and forced-choice odor).

By independently manipulating delay and size across a series of block trials, we would expect to see rats adapt their behavior accordingly to bias responses towards higher value outcomes (i.e. shorter delays and larger rewards). Across many studies, including current ones completed in this dissertation, we do indeed see rats change behavior to favor the higher value outcomes across many behavioral measures of the task (Roesch et al., 2009; Roesch and Bryden, 2011; Burton et al., 2014b; Hernandez et al., 2015). Specifically, the behavioral measures looked at are percent choice for free-choice trials, percent correct on forced-choice trials, and reaction-times for both forced- and free-choice trials. Percent choice was calculated based on the number of fluid well entries for each condition in each block for free-choice trials, for example how many times the animal entered the left well during a free-choice odor divided by the total number of free-choice trials in that block. Percent correct was calculated based on the number of correct fluid well entries for each forced-choice trial and each condition (short, long, big, and small). Reaction times were calculated based on odor cue offset to when the animal left the odor port for both forced- and free-choice trials for each condition (short, long, big, and small).

Predictably, we have shown that rats chose the high-value outcome significantly more often on free-choice trials (Figure 2B) and are better (Figure 2C) and faster (Figure 2D) on high-value forced-choice trials (Burton et al., 2014b). By analyzing behavior at the beginning and end of each block of trials we can reliably show changes in behavior mentioned above as the block progresses and the animal learns contingencies and biases behavior towards higher valued outcomes. The goal of my research was to investigate these behavioral measures after manipulations via

previous cocaine exposure or pharmacological lesion during performance of this task. After recording from single neurons while rats perform this task, I analyzed neuronal data by comparing firing rates in both response directions (contralateral or ipsilateral relative to recording electrode placement) for each trial type (forced- or free-choice) for each condition (short, long, big, or small) at the time point in which the animal is processing the odor cue and making a decision (i.e. odor onset to fluid well entry or odor onset to odor port exit, "odor epoch").

The striatum is necessary for proper performance on delay discounting tasks and it has also been shown that cocaine impacts performance by increasing impulsive choice in these tasks (Roesch et al., 2007b; Simon et al., 2007; Dandy and Gatch, 2009; Setlow et al., 2009; Mendez et al., 2010; Mitchell et al., 2014). Below is a description of the evidence that previous cocaine self-administration biases rat's choices toward higher value rewards on the odor-guided decision-making task and that encoding in striatal sub-regions was impacted by cocaine exposure. Furthermore, I show evidence that disruption of NAc in the striatal circuit can impact behavior and encoding in DLS that are important when considering the impact of drug exposure within this circuit.

Chapter 2: Previous cocaine self-administration disrupts reward expectancy encoding in ventral striatum

Authors: Amanda C. Burton[1,2], Gregory B. Bissonette[1,2], Daniela Vazquez[1], Elyse M. Blume[1], Maria Donnelly[1], Kendall C. Heatley[1], Abhishek Hinduja[1], and Matthew R. Roesch[1,2]

1. Department of Psychology
2. Program in Neuroscience and Cognitive Science
1147 Biology-Psychology Building
University of Maryland
College Park, MD 20742

ABSTRACT

The nucleus accumbens core (NAc) is a critical region in the basal ganglia that is important for integrating and providing information to downstream areas about the timing and value of anticipated reward. NAc is one of the first brain regions to be affected by drugs of abuse, yet we still do not understand how neural correlates related to reward expectancy are affected by previous cocaine self-administration. To address this issue, we recorded from single neurons in the rat NAc after previous sucrose (control) or cocaine self-administration during performance of an odor-guided decision-making task where we independently manipulated value of expected reward by changing the delay to or size of reward across a series of blocks of trials. We found that previous cocaine self-administration made rats more impulsive, biasing choice behavior toward more immediate reward. Further, compared to controls, cocaine-exposed rats showed significantly fewer neurons in the NAc that were responsive during odor cues and reward delivery, and in those reward-responsive neurons that remained, diminished directional and value encoding was observed. Lastly, we found that after cocaine exposure, reward-related firing during longer delays was reduced compared to controls. These results demonstrate that prior cocaine self-administration alters reward-expectancy encoding in NAc, which would contribute to poor decision-making observed after chronic cocaine use.

INTRODUCTION

The nucleus accumbens core (NAc) region of the ventral striatum (VS) is an important nucleus of the basal ganglia where relevant limbic and motor information is integrated to guide reward-based learning and decision-making (Mogenson et al., 1980; Roesch and Bryden, 2011; Burton et al., 2015). One mechanism that NAc uses to serve as this function is to encode the value of expected outcomes after decisions. It has been shown that many neurons in NAc fire after an instrumental response or presentation of a Pavlovian cue as animals anticipate the imminent delivery of reward (Bowman et al., 1996; Hassani et al., 2001; Setlow et al., 2003; Cromwell and Schultz, 2003; Janak et al., 2004; Nicola, 2007; Khamassi et al., 2008; Van der Meer and Redish, 2009; Ito and Doya, 2009; Roesch et al., 2009; Cai et al., 2011; Jimura et al., 2013; Tanaka et al., 2016). This signal is thought to play a critical role in timing rewarded outcomes during behavior (Meck et al., 2008; Singh et al., 2011; Mello et al., 2015; Takahashi et al., 2016). Consistent with this hypothesis, several studies have shown that loss of NAc is detrimental to behavior on temporal discounting tasks (Cousins et al., 1996; Cardinal et al., 2001, 2004; Bezzina et al., 2007; Floresco et al., 2008; Kalenscher and Pennartz, 2008; McDannald et al., 2011; Burton et al., 2014a) and alters neural signaling in downstream regions such as the VTA and dorsal striatum (DS) during similar tasks (Burton et al., 2014a; Takahashi et al., 2016). More specifically, the loss of NAc signaling eliminates reward prediction errors related to delayed outcomes encoded by dopamine (DA) neurons in the ventral tegmental area (VTA) (Takahashi et al., 2016), preventing DA cue selectivity from developing

(Takahashi and Schoenbaum, 2016), which would influence signaling in further downstream areas such as DS (Burton et al., 2014a).

Taken together, all of these studies suggest that the NAc is critical in integrating and providing information to downstream areas about the timing and value of obtained and expected reward to guide normal decision-making, a function that we know to be altered by chronic cocaine exposure; rats exposed to cocaine are more sensitive to delays to reward (Roesch et al., 2007b; Simon et al., 2007; Dandy and Gatch, 2009; Setlow et al., 2009; Mendez et al., 2010). This should come as no surprise because NAc is one of the first regions targeted by drugs of abuse due to its strong dopaminergic innervation (Koob and Volkow, 2010; Steinberg et al., 2014; Keiflin and Janak, 2015; Lüscher, 2016). Furthermore, any alterations of NAc would have profound impact on behavior due its influence on the motor system and dorsal striatum via spiraling connectivity with the DA system (Haber et al., 2000; Belin and Everitt, 2008; Haber and Knutson, 2009; Everitt and Robbins, 2013; Burton et al., 2015).

Thus, it is clear that NAc signals reward expectancies that are critical for timing of outcomes and normal delay discounting behavior, and that cocaine exposure impacts NAc and makes rats more impulsive. Yet, we still do not know the long-term impact that cocaine self-administration has on neural correlates in the NAc during complex decision-making tasks that vary reward expectancies. To address this issue, we recorded from single neurons in the rat NAc after previous sucrose (control) or cocaine self-administration

during performance of an odor-guided decision-making task where we independently manipulated value of expected reward by changing the delay to or size of reward across a series of blocks of trials. We found that previous cocaine exposure increased performance on free-choices toward immediate reward and decreased performance on forced-choice trials compared to controls. In addition, cocaine-exposed rats showed decreased numbers of neurons responsive to odor cues and rewards, and those that remained responsive exhibited diminished value and directional encoding within the NAc at the time of reward compared to controls.

MATERIALS AND METHODS

Subjects. Male Long-Evans rats were obtained at 175-200g from Charles River Laboratories. Rats were tested at the University of Maryland, College Park in accordance with UMD and NIH guidelines.

Experimental Timeline. The following is a general timeline of the experiment; more detailed explanations for each part of the experiment are outlined below. Rats were obtained at six weeks of age/average weight of 175-200g and acclimated to our facility for one week prior to behavioral training. Rats were trained on an odor-guided decision-making task in custom behavioral boxes for approximately six weeks prior to surgical implantation of a catheter in the jugular vein and a chronic advancing electrode just dorsal to the nucleus accumbens core region of ventral striatum (VS). Rats were able to recover for two weeks following surgery. Cocaine/sucrose self-administration was then implemented for 12 consecutive days in Med Associates behavioral boxes followed by a month-long withdrawal period. Single-unit neural

recordings (Plexon) during performance of the previously trained behavioral odor-guided decision-making task. Rats were transcardially perfused and standard histological techniques were implemented to ensure proper placement of electrodes.

Odor-guided decision-making task. Before surgery, all rats were trained on the odor-guided delay/size choice task for approximately six weeks. Rats were mildly water-deprived to ensure motivation to complete a session of task performance. Rats were shaped to nose-poke in a central odor panel that was controlled via computer to lengthen the amount of time spent in the odor port and the adjacent fluid wells to receive reward. Once proper responding was attained, rats were introduced to the different instructive odors. On each trial, nose poke into the odor port after house light illumination resulted in delivery of an odor cue to a hemicylinder located behind this opening. One of three different odors (2-Octanol, Pentyl Acetate, or Carvone) was delivered to the port on each trial. One odor instructed the rat to go to the left fluid well to receive reward (forced-choice), a second odor instructed the rat to go to the right fluid well to receive reward (forced-choice), and a third odor indicated that the rat could obtain reward at either well (free-choice). Odors were counterbalanced across rats. The meaning of each odor did not change across sessions. Odors were presented in a pseudorandom sequence such that the free-choice odor was presented on 7/20 trials and the left/right odors were presented in equal proportions.

During training and recording, one well was randomly designated as short (500 ms) and the other long (1-7s) at the start of the session (Figure 1A: Block 1). In the second block of trials, these contingencies were switched (Figure 1A: Block 2). The length of the delay under long conditions abided by the following algorithm: the side designated as long started as 1s and increased by 1s every time that side was chosen on a free-choice odor (up to a maximum of 7s). If the rat chose the side designated as long less than 8 out of the previous 10 free choice trials, the delay was reduced by 1s for each trial to a minimum of 3s. The reward delay for long forced-choice trials was yoked to the delay in free-choice trials during these blocks. In later blocks, we held the delay preceding reward delivery constant (500ms) while manipulating the size of the expected reward (Figure 1A: Blocks 3 and 4). The reward was a 0.05 ml bolus of 10% sucrose solution. For big reward, an additional one to two boli were delivered 500ms after the first bolus. At least 60 trials per block were collected for each neuron and session. Essentially there were four basic trial types (short, long, big, and small) by two directions (left and right) by two stimulus types (free- and forced-choice odor).

Surgeries and Recovery. All rats were implanted with catheters for self-administration and electrodes for single-unit recordings (sucrose control group, n = 6 and cocaine group, n =4). Rats were catheterized in the jugular vein with Silastic tubing (0.02x0.037in., by Dow Corning) with a modified 22G 5up cannula (Plastics One) on the end, which was then fed through the fascia layer over the shoulder and back and cemented to the skull next to the electrode implant site. Dummy blockers (Plastics One) were screwed on to the cemented modified cannula end to keep sterile

controlled access to catheters. For electrode implantation, rats were placed in a stereotaxic device and a drivable bundle of 10-25 µm diameter FeNiCr wires were chronically implanted in the left or right hemisphere dorsal to the nucleus accumbens core region of ventral striatum (1.6mm anterior to bregma, + or - 1.5mm laterally, and 6mm ventral to the brain surface). Electrode wires were housed in 27G cannula. Immediately prior to implantation, the wires were freshly cut with surgical scissors to extend ~1 mm beyond the cannula and electroplated with platinum (H2PtCl6, Aldrich) to an impedance of ~300kOhms. At the end of surgery, all catheters were flushed with 0.1 ml of a 5mg/ml dosage of the antibiotic gentamicin and a pain medication was given with saline to control pain and replenish fluids lost during surgery. Catheters were flushed with 0.1 ml of gentamicin (5mg/ml) every other day during recovery and through the self-administration phase (see below) of the experiment to prevent infection and keep catheters patent. An oral antibiotic (Cephalexin, 250mg/5ml) was administered twice a day for one week following surgery to prevent infection of the electrode implant site.

Self-administration and withdrawal. After rats recovered from surgery, a twelve-day self-administration protocol was implemented using Med Associates Inc. operant behavioral boxes. During days 1-6, the cocaine group (n = 4) self-administered 1mg/kg dosage of cocaine via lever press with a maximum of 30 infusions or 3-hour time limit. During days 7-12 a 0.5mg/kg dosage of cocaine was self-administered with a maximum of 60 infusions or a 3-hour time limit. Sessions began with illumination of house lights, extension

of the lever, and an initial illumination of a cue light above the lever for 2.3 seconds. Active lever presses were paired with the cue light above the lever, which stayed on for the duration of the cocaine infusion (2.3 seconds per active infusion). Active lever presses for drug infusion could only take place 20 seconds after the previous active lever press. Lever presses during the inactive period resulted in no cue-light illumination or reward delivery. The active lever was extended for the duration of the session. Rats were taken out of the operant boxes after the maximum infusion or maximum time limit was reached as described above.

The control group (n = 6) followed the same protocol as above, however during days 1-6 two sucrose pellets were delivered to a food cup upon active lever presses and during days 7-12 one sucrose pellet was delivered to a food cup upon active lever presses. All other parameters remained the same as the cocaine group: start of a session, maximum number of active lever presses, maximum time in the box, and activation and duration of cue lights above levers.

Following the twelve days of self-administration, both cocaine and control rats went through a month-long period of withdrawal in their home cages. Rats were not placed in any behavioral boxes during this time period. Their health was monitored closely and they were handled often each week during the withdrawal period.

Single-Unit Recording. Following the month-long withdrawal period, rats were placed in Plexon equipped behavioral boxes to perform the previously trained odor-guided choice task. Recording procedures were the same as described previously (Bryden et al., 2011b, 2011c) and as follows. Wires were screened for

activity daily; if no activity was detected, the rat was removed, and the electrode assembly was advanced 40 or 80μm. Otherwise active wires were selected to be recorded, a session was conducted, and the electrode was advanced at the end of the session. Neural activity was recorded using two identical Plexon Multichannel Acquisition Processor systems (Dallas, TX), interfaced with odor discrimination training chambers. Signals from the electrode wires were amplified 20X by an op-amp headstage (Plexon Inc, HST/8o50-G20-GR), located on the electrode array. Immediately outside the training chamber, the signals were passed through a differential pre-amplifier (Plexon Inc, PBX2/16sp-r-G50/16fp-G50), where the single unit signals were amplified 50X and filtered at 150–9000 Hz. The single unit signals were then sent to the Multichannel Acquisition Processor box, where they were further filtered at 250–8000 Hz, digitized at 40 kHz and amplified at 1–32X. Waveforms (>2.5:1 signal-to-noise) were extracted from active channels and recorded to disk by an associated workstation with event timestamps from the behavior computer. Waveforms were not inverted before data analysis.

Behavioral Analysis. Behavior during self-administration was evaluated by computing the average number of lever presses daily across rats in each group (cocaine and control) and computing the average number of lever presses across days 1-6 and days 7-12 across rats in each group (Figure 1F). Behavior during performance of the behavioral task was evaluated by computing the percent choice of value conditions (short, long, big, small) on free-choice trials as well as percent correct and reaction time (odor offset to

odor port exit) on forced-choice trials for each value condition (short, long, big, small) for the first ten and last ten trials of each block of trials (total of four blocks per session). Reaction time on free-choice trials was computed for all free-choice trials for each block. Percent choice was computed by determining how often the rat chose the more and less valued option on free-choice trials. Percent correct was computed by determining if the direction the animal chose was the same that was instructed by the forced-choice odor. Each of these behavioral analyses were computed and averaged across sessions for each group. Multi-factor analysis of variance (ANOVA) and post-hoc t-tests ($p < 0.05$) were used to determine differences in behavior between cocaine and control groups.

Neural Analysis. Recorded waveforms (>2.5:1 signal-to-noise) were extracted from active channels during recording sessions and recorded to disk by an associated workstation with event time-stamps from the behavioral computer. Extracted single-units were then sorted in Offline Sorter using template matching (Plexon) and exported to NeuroExplorer to determine time-stamped events related to spike activity. All further analysis was done using Matlab and Microsoft Excel. Analysis epochs were computed by taking the total number of spikes and dividing by time. Three analysis epochs were examined. One was taken 250ms before reward delivery to 1s after reward delivery (reward epoch) to capture activity related to expectancy and reward delivery. The 250ms period was chosen because it does not impinge on early activity after the response even at the shortest delay of 500ms. Another analysis was taken 100ms after odor onset to fluid-well entry (odor epoch). This period of time did not overlap with the reward epoch and captured activity before entering the fluid-well

delay period. Both epochs were compared to baseline (1 second before odor onset; Wilcoxon; $p < 0.05$) to determine task responsiveness. Lastly, for analysis of long delays we also used an epoch starting 500ms after fluid well entry to capture initial bursts of activity that were elicited upon well entry. This epoch did not overlap with the reward epoch even at the shortest long-delay trial (1s). The average population firing rates for control and cocaine-exposed rats were determined by determining each neuron's preferred and non-preferred direction (neural response that elicited the most activity (spikes/second)), then plotting normalized spike activity during the reward epoch aligned to reward delivery or well entry.

Wilcoxon rank sum ($p < 0.05$) analyses were used to determine differences in firing between preferred and non-preferred direction and for each value outcome for both controls and cocaine rats during odor and reward epochs. Chi-squares ($p < 0.05$) were performed on counts of neurons in each group to determine if there were any significant differences between groups. For each neuron, difference scores were computed between high and low value rewards (short-long/short+long and big-small/big+small). These index frequencies were then plotted in histograms and Wilcoxon rank sum ($p < 0.05$) tests were used to measure significant shifts from zero for each value distribution and to determine if these shifts were significantly stronger between groups.

RESULTS

Self-administration

All rats were trained on the reward-guided decision-making task (Figure 1A) prior to implantation of electrodes in VS (Figure 1G-H) and catheters for cocaine self-administration (see methods for more detail). During performance of the reward-guided decision-making task, on each trial, rats responded to one of two adjacent wells after sampling an odor at a central port (Figure 1A). Rats were trained to respond to three different odor cues: one odor that signaled reward in the right well (forced-choice), a second odor that signaled reward in the left well (forced-choice), and a third odor that signaled reward at either well (free-choice). Across blocks of trials in each recording session, we manipulated either the length of the delay preceding reward delivery (Figure 1A; Blocks 1-2; ~60 trial/block) or the size of the reward (Figure 1A; Blocks 3-4; ~60 trial/block).

After surgery and recovery, rats self-administered sucrose pellets or cocaine paired with a cue light over the course of twelve days. During days 1-6 (1 mg/kg cocaine or 2 sucrose pellets), the average number of active lever presses across rats out of a maximum of 30 was 20.2 (±9.8; standard deviation; s.d.) and 30 (±0 s.d.) for cocaine and sucrose, respectively (Figure 1F). During days 7-12 (0.5 mg/kg cocaine or 1 sucrose pellet), the average number active lever presses across rats out of a maximum of 60 was 51 (±14.1 s.d.) and 60 (±0 s.d.) for cocaine and sucrose, respectively (Figure 1F).

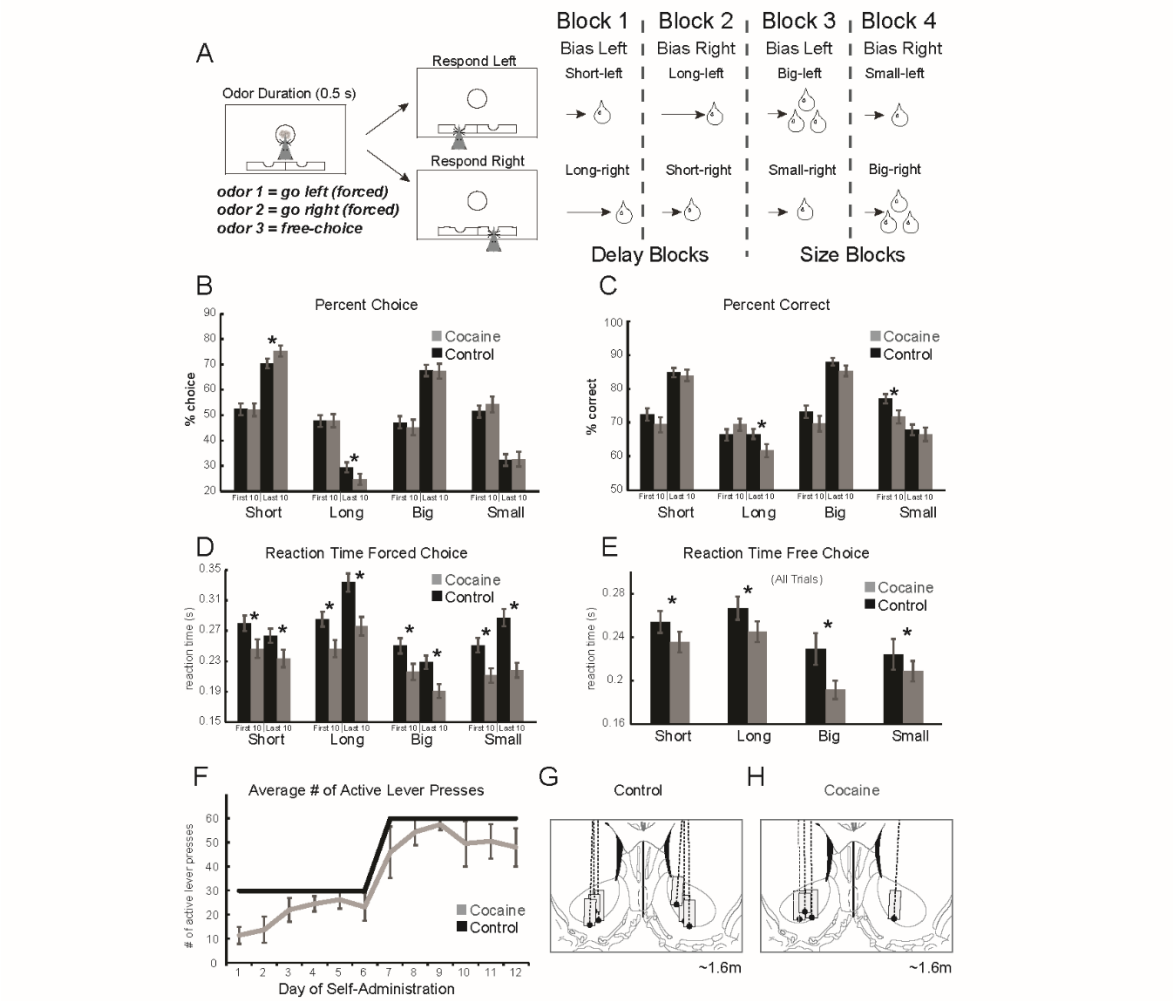

Figure 1: **A.** Task schematic, showing sequence of events in one trial (left panels) and the sequence of blocks in a session (right). Rats were required to nose-poke in the odor port for 0.5 s before the odor turned on for 0.5 s instructing them to respond to the adjacent fluid wells below where they would receive liquid sucrose reward after 500-7000 ms. For each recording session, one fluid well was arbitrarily designated as short (a short 500-ms delay before reward) and the other designated as long (a relatively long 1- to 7-s delay before reward) (Block 1). After the first block of trials (~60 trials), contingencies unexpectedly reversed (Block 2). With the transition to Block 3, the delays to reward were held constant across wells (500ms), but the size of the reward was manipulated. The well designated as long during the previous block now offered 2–3 fluid boli, whereas the opposite well offered 1 bolus. The reward stipulations again reversed in Block 4. **B.** Percent choice on free-choice trials in each value manipulation over the first ten and last ten trials of each block averaged across animals and sessions (controls, black bars; cocaine, gray bars). **C.** Percent correct on forced-choice trials in the same manner as B. **D.** Reaction time (odor port exit minus odor offset) on forced-choice trials in the same manner as B and C. **E.** Reaction time (odor port exit minus odor offset) on all free-choice trials for each value manipulation. For these analyses, behavior was looked at by session. **F.** Average number of active lever presses during sucrose (black) or cocaine (gray) self-administration averaged across rats for each day (days 1-12). **G-H.** Location of recording sites (Paxinos and Watson). Gray boxes mark the extent of the recording locations. Error bars indicate SEM. Asterisks (*) indicate significance ($p < 0.05$) in multi-factor ANOVA and/or post-hoc t-tests.

Cocaine made rats more sensitive to delay and size manipulations

After a month-long withdrawal period, we recorded from single units in the nucleus accumbens core (NAc) from six sucrose control rats and four cocaine-exposed rats during performance of the odor-guided decision-making task. Behavior was looked at by session, with a total of 154 sessions from controls (20, 22, 21, 29, 25, and 37) and 95 sessions from cocaine rats (19, 25, 28, and 23). Multi-factor ANOVAs were performed to assess performance in the task for several behavioral measures including percent correct, percent choice, and reaction time on both forced- and free-choice trials (odor offset to odor port exit). Factors in the ANOVA included group (sucrose controls or cocaine), value (high or low), value manipulation (size or delay), and phase of learning (early: first ten trials or late: last ten trials).

Performance on free-choice trials was similar to previous results using the same task in that we see a bias toward higher value rewards (Roesch et al., 2007b). In the ANOVA with percent choice as the dependent variable and the factors described above there was a significant main effect of value ($F(1,1976) = 1025.4$, $p < 0.01$), as both control and cocaine rats chose short-delay and large-reward more often than long-delay and small-reward, respectively (Figure 1B). There was also a significant interaction between group, value and value manipulation ($F(1,1976) = 1592.6$, $p < 0.01$) and between group, value and phase ($F(1,1976) = 1398.4$, $p < 0.01$) with cocaine rats choosing the high-value reward significantly more often than controls in the last ten free-choice trials during delay manipulations (Figure 1B; $t(247) = 2.86$, $p < 0.01$). As shown previously, cocaine-exposed rats were strongly drawn to the more immediate reward at the end of delay blocks (Roesch et al., 2007b; Simon et al.,

2007; Dandy and Gatch, 2009; Setlow et al., 2009; Mendez et al., 2010). Additionally, cocaine rats were significantly faster on all free-choice trial types compared to controls (Figure 1E; ANOVA; main effect of group ($F(1,988) = 6.03$, $p < 0.05$). Although cocaine made rats more sensitive to delay manipulations, cocaine and control rats chose large over small reward at similar rates (Figure 1B).

In separate ANOVAs with percent correct and reaction time on forced-choice trials as the dependent variables and the same factors as stated above, we found a main effect of value (percent correct: $F(1,1976) = 201.66$, $p < 0.01$; forced-choice reaction time: $F(1,1976) = 29.01$, $p < 0.01$) and an interaction of value and phase (percent correct: $F(1,1976) = 199.08$, $p < 0.01$; forced-choice reaction time: $F(1,1976) = 28.41$, $p < 0.01$). This indicates that overall, both control and cocaine rats were significantly better and faster on high value forced-choice trial outcomes particularly in the late phase of each block (Figure 1C-D). However, there was also a main effect of group in the ANOVAs on forced-choice behavioral measures (percent correct: $F(1,1976) = 10.65$, $p < 0.01$; forced-choice reaction time: $F(1,1976) = 83.91$, $p < 0.01$), with cocaine rats being significantly faster and worse on forced-choice trials compared to controls (Figure 1C-D). Although both groups were able to track value across blocks in the task, previous cocaine exposure did impair general performance on forced-choice trials.

We conclude that previous cocaine self-administration had a long-term impact on behavior during performance of the odor-guided decision-making

task. Overall, cocaine rats exhibited stronger response biases toward more immediate reward on free-choice trials towards the end of delay blocks. In addition, cocaine rats were also significantly faster and worse on forced-choice trials overall. These results are consistent with previous work demonstrating that cocaine self-administration makes rats more impulsive (Roesch et al., 2007b; Simon et al., 2007; Dandy and Gatch, 2009; Setlow et al., 2009; Mendez et al., 2010).

Previous cocaine self-administration alters reward-related activity in NAc compared to controls

In control and cocaine-exposed rats, 195 (30%) and 103 (21%) neurons increased firing during the reward epoch (250ms before reward delivery to 1s after reward delivery) compared to baseline (1 second before odor onset; Wilcoxon, $p < 0.05$), respectively. The frequency of responsive neurons was significantly different between groups, with more increasing-type neurons in NAc in controls compared to cocaine rats (chi-square = 6.47, $p < 0.05$). In Figure 2A, we show a single-unit example showing anticipatory activity during performance of a long delay trial where firing to delayed rewards remains elevated until reward delivery (left panel) and a sharp and stronger increase in activity when reward was delivered after a short delay (right panel).

To further look at the effects of reward on neural activity in the NAc, next we plotted the normalized average population activity for the last ten

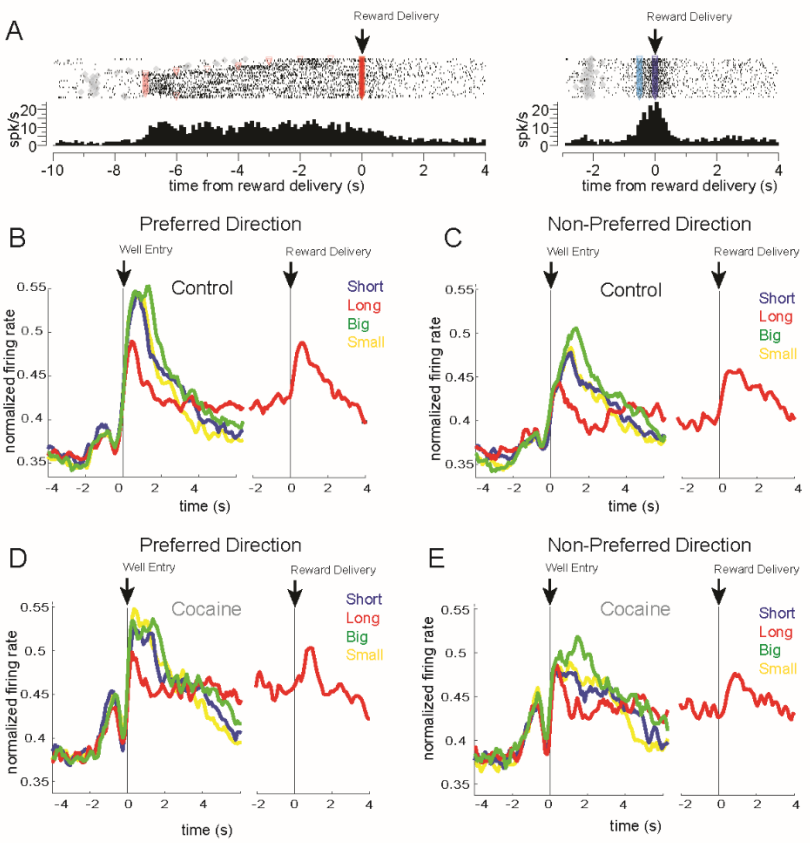

Figure 2: Previous cocaine-exposure diminishes value and directional encoding in NAc. **A.** Single neuron example during long-delay (left panel) and short-delay trials (right panel). Activity is aligned to reward delivery (marked by arrows) and binned at 100ms. 1 tick mark equals 1 action potential. **B-C.** Population activity for increasing-type neurons in control rats (n = 195) during the last ten trials for each outcome. B shows normalized firing rates for increasing-type neurons in the preferred direction (direction that elicited the most activity). C shows normalized firing rates for increasing-type neurons in the non-preferred direction). Blue lines are short-delay trials, red lines are long-delay trials, green lines are large-reward trials, and yellow lines are small-reward trials). Activity is aligned to fluid-well entry, and for long-delay trials we split activity and aligned on reward-delivery (right panels in B and C) to allow observation of reward delivery after longer delays. **D-E.** Population activity for increasing-type neurons in cocaine rats (n = 103) during the last ten trials for each outcome. Activity is plotted in the same manner as B-C.

trials for all increasing neurons for controls and cocaine rats aligned to fluid-well entry for all trial types and reward-delivery on long-delay trials (Figure 2B-E). To make these plots, we determined which direction elicited the most activity and designated this as the preferred direction (Figure 2B and 2D, left panels) and the other direction as the non-preferred direction (Figure 2C and 2E, right panels). In these plots, the colored lines reflect the neural response to the value of reward on that particular set of trials and the delay is split on long delay trials so that activity can be aligned to both fluid-well entry and reward delivery (for other trial-types the delay was a fixed 500 ms). 'Blue' lines represent activity on short-delay trials, 'red' lines represent activity on long-delay trials, 'green' lines represent activity on large-reward trials, and 'yellow' lines represent activity on small-reward trials.

As expected in controls, activity in NAc in the preferred direction was modulated by value; there was an increase in firing for high-value outcomes compared to low-value outcomes (Figure 2B; blue (short) and green (big) lines higher than red (long) and yellow (small) lines, respectively). In cocaine rats this value-encoding was diminished in the preferred direction (Figure 2D). To quantify this effect, we computed value indices on firing rates during the reward epoch by taking difference scores for each increasing-type neuron for each value manipulation (short-long/short+long and big-small/big+small). These value index frequencies were then plotted in histograms and Wilcoxon rank sum ($p < 0.05$) tests were used to measure significant shifts from zero for each value distribution within each group and to determine if distribution were significantly different between groups (Figure 3). Black bars represent indices

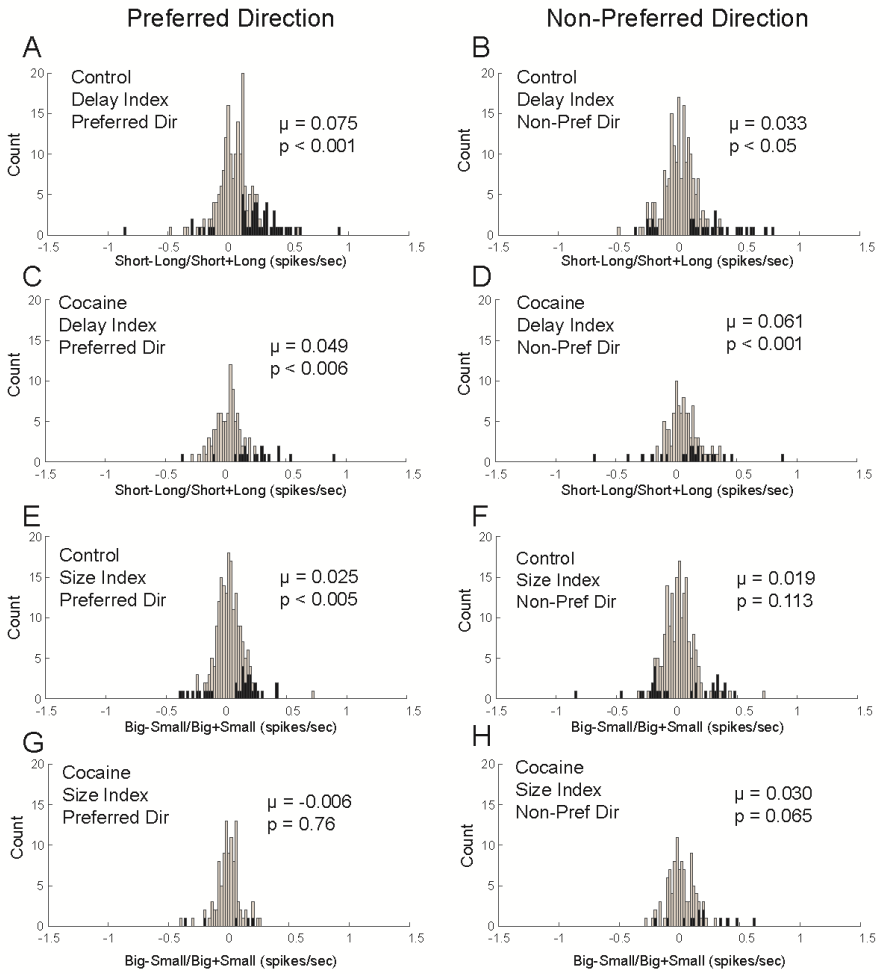

Figure 3: Distribution of value indices for delay (**A-D**) and size manipulations (**E-H**) for control (A-B, E-F) and cocaine rats (C-D, G-H) in the preferred (left panels) and non-preferred direction (right panels). Delay index = normalized firing rates for short-long/short+long during the reward-epoch; size index = normalized firing rates for big-small/big+small during the reward epoch. Neural activity was taken 250ms before reward delivery to 1s after reward delivery. Black bars represent neurons that showed significant modulation of expected outcome (Wilcoxon, $p < 0.05$).

from neurons that exhibited significant firing rated difference between differently valued rewards at the single unit level (i.e., short vs long; big vs small; Wilcoxon; $p < 0.05$).

We found that control rats showed a significant positive shift from zero for both the delay (Figure 3A, $p < 0.01$, $\mu = 0.075$) and size index (Figure 3E, $p < 0.01$, $\mu = 0.025$), indicating that the majority of NAc neurons in control rats exhibited significantly increased activity for higher value outcomes (short and larger rewards, respectively) in the preferred direction. Like controls, in cocaine rats there was also a significant positive shift in the distribution of delay indices in the preferred direction (Figure 3C, $p < 0.01$, $\mu = 0.049$), however, unlike controls, there were no significant shifts from zero in the distribution of size indices (Figure 3G, $p = 0.76$, $\mu = -0.0065$), indicating that NAc neurons did not tend to fire more strongly for large compared to small reward after cocaine self-administration.

When directly comparing the index distributions between control and cocaine rats, we found that there were significant differences between groups (Wilcoxon rank sum, $p < 0.05$). The shifts above zero as discussed above in both the delay and size indices in the preferred direction for control rats was significantly stronger compared to cocaine rats ($p < 0.05$), indicating stronger value encoding of higher value outcomes (short delays and larger rewards) in the NAc of control rats during both delay and size manipulations.

Comparing counts of single neurons that were significantly selective for size and delay further supports the finding that value selectivity was reduced after cocaine self-administration. To illustrate this, we performed a Wilcoxon rank sum test ($p <$

0.05) on firing rates for each individual neuron in the preferred and non-preferred direction and for each value condition in the task (short, long, big, small; black bars in Figures 3A-H). When looking at each value condition in the preferred direction, control rats showed significant differences between counts of increasing neurons for short compared to long (short n = 43, long n = 8, Figure 3A, chi-square = 23.88, $p < 0.01$) and between counts of increasing neurons for big compared to small (big n = 23, small n = 10, Figure 3E, chi-square = 5.04, $p < 0.05$). Cocaine rats showed a similar pattern between counts of increasing neurons in the preferred direction for short compared to long (short n = 16, long n = 2, Figure 3C, chi-square = 8.76, $p < 0.01$). However, there were minimal counts of neurons in the preferred direction for big and small conditions and the difference between counts was not significant in cocaine rats (big n = 3, small n = 2, Figure 3G, chi-square = 0.162, p = .687).

Overall, neural activity in the preferred direction in the NAc at the time of reward in control rats seems to be encoding the value of the expected reward in a value-dependent manner. We see increased activity and increased counts of neurons for short compared to long rewards and for big compared to small rewards in the preferred direction in controls. However, in cocaine rats this was diminished, particularly for size manipulations. Next, we consider firing in the non-preferred direction. Unlike the preferred direction, weaker value encoding was observed in control rats in the non-preferred direction for both delay (although still significant; Figure 3B, $p < 0.05$, $\mu = 0.033$) and size

blocks (Figure 3F, $p = 0.113$, $\mu = 0.019$). In rats which had been exposed to cocaine, the opposite was true: cocaine rats showed significantly stronger value encoding in the non-preferred direction on delay blocks compared to the preferred direction (Figure 3D, $p < 0.05$, $\mu = 0.061$) and this was significantly stronger compared to controls (Wilcoxon rank sum, $p < 0.05$). In addition, cocaine rats showed stronger value encoding that approached significance in the non-preferred direction on size blocks compared to the preferred direction (Figure 3H, $p = 0.06$, $\mu = 0.030$). Thus, value encoding was stronger in the preferred direction compared to the non-preferred direction for both delay and size manipulations in controls and this was not true for the cocaine group. In cocaine animals, value encoding was weaker in the preferred direction compared to the non-preferred direction on delay blocks, and was not significantly different on size blocks for either direction.

We conclude that neural activity in the NAc at the time of reward in the preferred direction is strongly encoding value in control rats and that cocaine exposure diminishes this functionality. Further, value encoding is overall less directional in cocaine rats, in that reward selectivity was similarly encoded in both response directions. Lack of directional selectivity in NAc of cocaine-exposed rats is further demonstrated by simply comparing the counts of neurons whose activity – averaged across outcomes – was significantly different between directions (Wilcoxon; $p < 0.05$). Rats in the cocaine groups had significantly fewer direction-selective neurons during the reward epoch compared to controls (controls: $n = 53$ (27%), cocaine: $n = 14$ (14%) chi-square $= 4.08$, $p < 0.05$).

Firing on longer delays was attenuated after cocaine self-administration

In the sections above we demonstrate weaker encoding of size and delay at both the population and single neuron level after cocaine exposure. Overall, cocaine rats had fewer neurons that increased firing in anticipation and delivery of reward, and those that were responsive exhibited weaker encoding. Here we ask how firing rates on trials with delayed reward changed as delays got longer. In particular, we were interested in knowing if selectivity in cocaine rats was overly sensitive to longer delays. To address this issue, we plotted firing rates over all neurons that increased firing in both controls and cocaine rats during reward delivery, broken down by delays 1-5 seconds, aligned to fluid well entry (six and seven second delays were excluded due to insufficient numbers of trials). We examined the reward epoch as above, but also an earlier epoch – 500 ms after well entry – that encompassed time immediately upon entry in the fluid well at the start of each delay, without impinging on firing to actual reward delivery (i.e., during the reward epoch). For both epochs, we found that firing declined as delays became longer, this is illustrated in Figure 4A-B, which represents firing aligned to well entry for delays 1-5 seconds. Although this decrement was similar for 1-3 second delays in both groups, firing stabilized in control rats during longer delays (Figure 4A), but continued to decline in rats that had self-administered cocaine (Figure 4B).

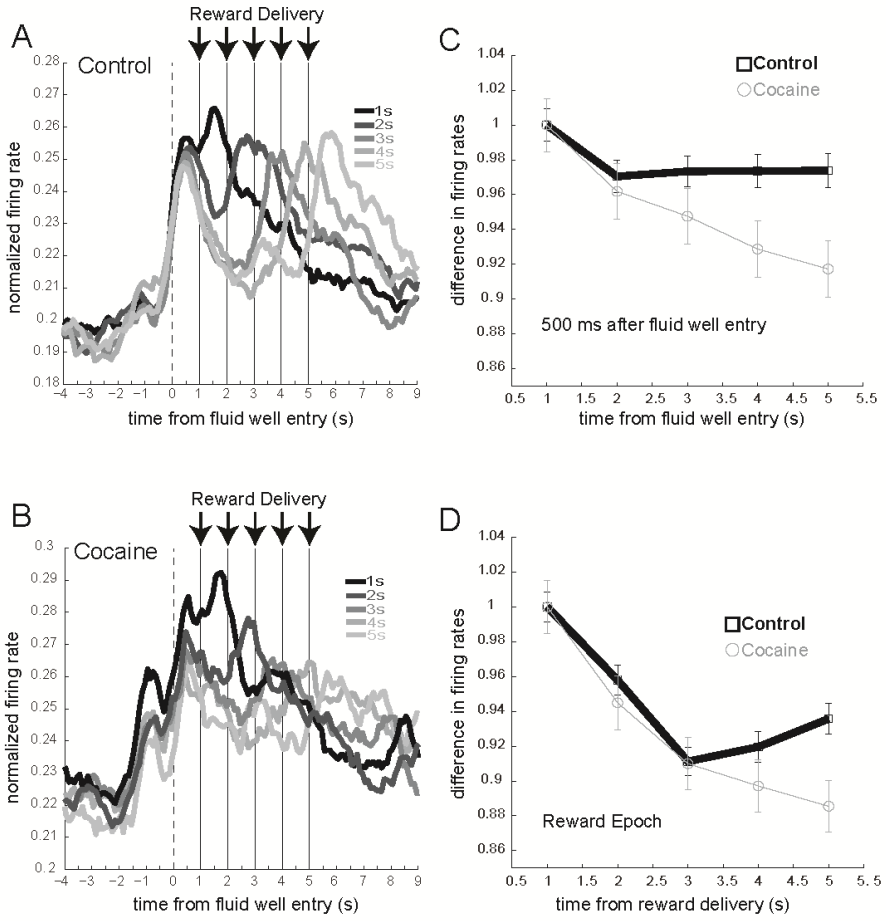

Figure 4: Firing on longer delays was attenuated in cocaine animals. **A.** Normalized average firing rates across increasing-type neurons during reward delivery broken down by delays 1-5 seconds for control rats. Activity is aligned to fluid well entry in order to encompass time immediately upon entry in the fluid well at the start of each delay before reward delivery. Reward delivery is marked by arrows for each delay. Delays 1-5 seconds are represented by black to lightest gray lines with 1 second delay being black lines and lightest gray lines being 5 second delays. **B.** Normalized average firing rates across increasing-type neurons during reward delivery broken down by delays 1-5 seconds for cocaine rats. Activity is plotted in the same manner as A. **C.** Normalized average firing rates, divided by maximum firing rate on 1s delays 500ms after fluid-well entry. Control rats are plotted in black squares/lines, cocaine rats are plotted in gray circles/lines. **D.** Normalized average firing rates, divided by maximum firing rate on 1s delays during reward delivery epoch (250ms before reward delivery to 1s after reward delivery). Control rats are plotted in black squares/lines, cocaine rats are plotted in gray circles/lines.

In both groups, firing increased at the time when rats entered the fluid well. Notably this initial response was similar across the 5 delays in controls, but continuously declined in rats that had self-administered cocaine. In controls during this period (500ms after fluid well entry), firing rates were not significantly different during delays 2-5 seconds compared to the 1 second delay (Figure 4C; Wilcoxons; p's > 0.05). Thus, in controls, after the initial decrease in firing when delays lengthened (i.e., 0.5s to 1s), firing at the beginning of the delay, upon well entry, remained stable. This was not true in rats exposed to cocaine in that firing continued to drop as delays become longer; firing was significantly reduced during delays 2-5 seconds compared to 1 second delay in the cocaine group (Figure 4C; Wilcoxon; $p < 0.05$).

Consistent with the main analysis described above both groups showed reduced firing rates for rewards at longer delays; firing to reward during 2-5 second delays was significantly reduced compared to firing during the 1 second delay in both cocaine and control rats (Figure 4D; Wilcoxon; $p < 0.05$). Also notable is that, during the 5 second delay, firing was significantly higher in controls compared to the cocaine groups (Wilcoxon; $p < 0.05$). Unfortunately, there were not enough trials to adequately determine if this finding persisted through 6 and 7 second delays, thus the significance of this result requires further study.

Reward selectivity during the odor epoch was not altered by cocaine

Above we showed that neural encoding during the reward epoch was affected by cocaine self-administration. Here, we perform the same analysis but during the period between odor onset and well entry (i.e., odor epoch = 100ms after odor onset to fluid well entry). In control and cocaine-exposed rats, 175 (27%) and 99 (20%) neurons increased firing during the odor cue period compared to baseline (1 second before odor onset; Wilcoxon, $p < 0.05$), respectively. The frequency of responsive neurons was significantly different between groups, with more increasing-type neurons in NAc during the odor epoch in controls compared to cocaine rats (controls n = 175, cocaine n = 99, chi-square = 3.92, $p < 0.05$). In these neurons, we found no significant differences between counts of neurons in controls and cocaine rats in the number of neurons that showed increases for one direction over another (Wilcoxon; $p < 0.05$; controls n = 18 (10%), cocaine n = 20 (20%), chi-square = 3.18, $p = 0.07$). Further, there were no differences in counts of neurons that encoded predicted outcomes (short, long, big, and small) in the preferred direction (Wilcoxon; $p < 0.05$) during delay blocks (controls: short n = 8, long n = 8; cocaine: short n = 13, long n = 3; chi-square = 2.22, $p = 0.13$) or size blocks (controls: big n = 20, small n = 10; cocaine: big n = 11, small n = 5; chi-square = 0.03, $p = 0.85$).

As above we computed value indices by taking difference scores for each increasing-type neuron in each value manipulation (short-long/short+long and big-small/big+small). When comparing the index distributions between control and cocaine rats, we found that there were no significant differences between groups (Wilcoxon rank sum, $p > 0.05$). There were no differences between controls and cocaine rats in index distributions for the delay index in the preferred ($p = 0.2466$)

and non-preferred direction ($p = 0.9605$). Even though cocaine rats showed a significant positive shift from zero in the preferred direction in the size index and controls did not, there was no significant difference in the strength of this shift between controls and cocaine rats ($p = 0.4970$) and no significant difference for the size index in the non-preferred direction ($p = 0.1046$).

Thus, overall, we conclude that although there were significantly fewer counts of neurons responsive prior to well entry during odor presentation and the decision to move, there were no significant differences between controls and cocaine rats in what was encoded during this period.

DISCUSSION

The nucleus accumbens core (NAc) is thought to be one of the first brain regions to be affected by drug abuse, with intense dopaminergic innervation from midbrain dopamine neurons that are initially targeted by drugs of abuse (Koob and Volkow, 2010; Keiflin and Janak, 2015; Lüscher, 2016). Previously, we and others have shown that lesions to NAc alter activity in downstream areas irrespective of drug use (Burton et al., 2014a; Takahashi et al., 2016). Here, we wanted to examine neural activity in the NAc after previous cocaine-exposure while rats performed a reward-guided decision-making task where the value of reward was independently manipulated by changing the delay to and size of reward. As described in several previous studies, rats previously exposed to cocaine more strongly biased behavior to rewards delivered after a shorter delay (Roesch et al., 2007b; Simon et al., 2007; Dandy and Gatch, 2009; Setlow et al., 2009; Mendez et al., 2010).

Consistent with cocaine's impact on behavior, we saw several alterations in NAc that might contribute to abnormal task performance. First, we observed an overall reduction in the number of cells responsive to both odor cues and reward delivery. Second, in those cells that remained reward-responsive, we observed a reduction in directional and value-encoding. Fewer neurons and weaker encoding during important time periods in the task might underlie drug-induced changes in behavior. Third, neural firing to delayed rewards was weaker in rats that had self-administered cocaine. This is of particular importance, as cocaine rats were more sensitive to longer delayed rewards in favor of the more immediate reward.

Although altered and lost directional and value encoding in the NAc could affect several downstream systems, we speculate that the behavioral and neurophysiological results described here are tightly linked to changes in dopamine (DA) signaling. The spiraling DA circuit between midbrain regions and striatal targets is known to be affected by drug-seeking and drug-taking (Haber et al., 2000; Belin and Everitt, 2008; Haber and Knutson, 2009; Everitt and Robbins, 2013). Reward prediction errors (reward experienced minus reward expected) are critical in learning about the current environment and are signaled by DA neurons (Schultz, 2006). It has recently been shown that reward prediction errors signaled by dopaminergic neurons in the VTA depend upon accurate signaling from the NAc, specifically in relation to the timing of rewarded outcomes (Takahashi et al., 2016) and that this is essential to developing neural selectivity to cues in DA neurons and neural signaling in the NAc (Nicola et al., 2005; Takahashi and Schoenbaum, 2016).

Several studies have also shown changes in encoding DA signaling in NAc even after periods of abstinence (Saddoris and Carelli, 2014; Cameron et al., 2016; Saddoris et al., 2016a). Specifically, cocaine-exposure has been shown to diminish neural activity in the NAc and DA release in the NAc to cues predicting rewards and punishments during reversal and Pavlovian discrimination tasks (Takahashi et al., 2007; Stalnaker et al., 2009; Saddoris et al., 2016b). Taken together, these findings fit with our overall result of fewer task responsive neurons in the NAc after cocaine exposure, which ultimately could impact processing cues, rewards, and associated actions in downstream areas such as the dorsal striatum (DS) as will be discussed in later chapters.

We conclude that previous cocaine self-administration has a long-term impact on decision-making and signaling in the NAc during performance of a complex behavioral decision-making task. Overall, we saw a behavioral bias toward shorter delays and a reduction in cells in the NAc that responded to odor cues and expected rewards after cocaine exposure. In addition, the reward-responsive cells that remained after cocaine exposure showed diminished value and directional selectivity. Reduced encoding of value and direction suggests that cocaine would impair the ability of NAc to guide behavior via model-based mechanisms that are reliant on this information to normally guide decision-making.

Chapter 3: Prior cocaine self-administration increases response-outcome encoding that is divorced from actions selected in dorsal lateral striatum

Published Manuscript

The Journal of Neuroscience, In Press

Authors: Amanda C. Burton[1,2], Gregory B. Bissonette[1], Adam C. Zhao[1], Pooja K. Patel[1] and Matthew R. Roesch[1,2]

1. Department of Psychology
2. Program in Neuroscience and Cognitive Science
1147 Biology-Psychology Building
University of Maryland
College Park, MD 20742

ABSTRACT

Dorsal lateral striatum (DLS) is a highly associative structure which encodes relationships between environmental stimuli, behavioral responses, and predicted outcomes. DLS is known to be a downstream target from the nucleus accumbens core (NAc), which shows diminished value and directional encoding after cocaine self-administration (Chapter 2). DLS is also known to be disrupted after chronic drug abuse, however it remains unclear what neural signals in DLS are altered. Current theory suggests that drug use enhances stimuli-response processing at the expense of response-outcome encoding, but this has mostly been tested in simple behavioral tasks. Here, we ask what neural correlates in DLS are affected by previous cocaine exposure as rats perform a complex reward-guided decision-making task where predicted reward value was independently manipulated by changing the delay to or size of reward associated with a response direction across a series of trial blocks. After cocaine self-administration, rats exhibited stronger biases toward higher value reward and firing in DLS more strongly represented action-outcome contingencies independent from actions subsequently taken rather than outcomes predicted by selected actions (chosen-outcome contingencies) and associations between stimuli and actions (stimulus-response contingencies). These results suggest that cocaine self-administration strengthens action-outcome encoding in rats (as opposed to chosen-outcome or stimulus-response encoding), which abnormally biases behavior toward valued reward when there is a choice between two options during reward-guided decision-making.

INTRODUCTION

Chronic drug use is thought to impair model-based goal-directed mechanisms governed by response-outcome (R-O) encoding while enhancing model-free stimulus-guided processing that controls habits via stimulus-response (S-R) encoding (Robbins and Everitt, 1999; Jentsch and Taylor, 1999; Everitt et al., 2001; Cardinal et al., 2002; Yin et al., 2004; Vanderschuren et al., 2005; Daw et al., 2005, 2011, Everitt and Robbins, 2005, 2016; Hyman et al., 2006; Belin and Everitt, 2008; Tricomi et al., 2009; Balleine and O'Doherty, 2010; Koob and Volkow, 2010; Lucantonio et al., 2012, 2014; Wied et al., 2013). Although these theories are well supported, few have actually recorded from the brains of cocaine-exposed animals to determine whether such correlates are altered. Further, the large majority of this theory is based on simple paradigms meant to isolate behaviors that are under the control of R-O and S-R associations in tasks where there is a singular response or no real choice at all (Pavlovian) (Robbins and Everitt, 1999; Everitt et al., 2001; Everitt and Robbins, 2005; Schoenbaum and Setlow, 2005; Vanderschuren et al., 2005; Everitt and Robbins, 2016; Nelson and Killcross, 2006, 2013; Nordquist et al., 2007; Ostlund and Balleine, 2008; Redish et al., 2008; Hogarth et al., 2013; LeBlanc et al., 2013; Corbit et al., 2014; Lucantonio et al., 2014; Schmitzer-Torbert et al., 2015). Although these studies are elegant in their design and have provided critical information to the field of drug abuse and neural control of behavior, work is necessary to understand how brain function changes in behaviors that offer choices between different rewards in situations where there is not necessarily a simple and direct pairing between stimuli, responses, and outcomes. Such paradigms might better reflect everyday decision-

making and are already known to evoke multiple neural representations outside the classic domain of R-O and S-R correlates that might be impacted by drugs of abuse.

Using this task, in the previous chapter we observed significantly fewer numbers of neurons in the nucleus accumbens core (NAc) responsive to both odor and reward epochs in the task compared to controls. In reward-responsive neurons that remained, we observed significantly diminished value and directional encoding in rats exposed to cocaine. Here, we examine firing in dorsal lateral striatum (DLS), an area known to be downstream from NAc and to be impacted by chronic cocaine use (Vanderschuren et al., 2005; Takahashi et al., 2007; Belin and Everitt, 2008; Everitt and Robbins, 2013; Lucantonio et al., 2014). In dorsal striatum, at least two different types of R-O correlates have been described in animals performing two-choice paradigms: one that reflects the relationship between *selected* actions and the outcomes that those actions predict – referred to as 'chosen-value' – and another that reflects the value of *possible* actions independent from the action that will ultimately be selected – referred to as 'action-value' (Lau and Glimcher, 2007, 2008; Nakamura et al., 2012). 'Action-value' signals are thought represent the relationship between potential actions and their predicted value so that they can be compared during decision making. Others refer to similar signals as a correlate of 'response-bias' (Lauwereyns et al., 2002; Nakamura et al., 2012). Response-bias signals in striatum emerge prior to the instruction to move, reflecting the behavioral bias the animal has for one direction over another.

Currently it is unknown how these neural representations are impacted by chronic cocaine use. However, we do know that when faced with a choice between

two options, rats that have been previously exposed to cocaine will more strongly bias their behavior toward more valuable options compared to controls (Roesch et al., 2007b; Simon et al., 2007; Mendez et al., 2010). To better understand the neural correlates that give rise to this behavioral effect, we implemented a two-week cocaine self-administration protocol with a one-month withdrawal period and then recorded from single neurons in DLS while rats performed a reward-guided decision-making task.

MATERIALS AND METHODS

Subjects. Male Long-Evans rats were obtained at 175-200g from Charles River Laboratories. Rats were tested at the University of Maryland, College Park in accordance with UMD and NIH guidelines.

Reward-guided decision-making. Before surgery, all rats were trained on the reward-guided decision-making task (Figure 1A), same as described above, for approximately six weeks.

During training and recording, one well was randomly designated as short (500 ms) and the other long (1-7s) at the start of the session (Figure 1A: Block 1). In the second block of trials, these contingencies were switched (Figure 1A: Block 2). The length of the delay under long conditions abided by the following algorithm: the side designated as long started off as 1s and increased by 1s every time that side was chosen on a free-choice odor (up to a maximum of 7s). If the rat chose the side designated as long less than 8 out of the previous 10 free choice trials, the delay was reduced by 1s for each trial to a minimum of 3s. The reward delay for long forced-

choice trials was yoked to the delay in free-choice trials during these blocks. In later blocks, we held the delay preceding reward delivery constant (500 ms) while manipulating the size of the expected reward (Figure 1A: Blocks 3 and 4). The reward was a 0.05 ml bolus of 10% sucrose solution. For big reward, an additional bolus was delivered 500 ms after the first bolus for all big-reward trials in that block. At least 60 trials per block were collected for each neuron. Essentially there were four basic trial types (short, long, big, and small) by two directions (left and right) by two stimulus types (free- and forced-choice odor).

Surgery. Seven rats were implanted with catheters for self-administration and electrodes for single-unit recordings in one survival surgery (cocaine group, n = 5 and control group, n = 2). Four additional control rats received electrodes only and were used as controls in another study (Chapter 4). Rats were catheterized in the jugular vein with Silastic tubing (0.02x0.037in., Dow Corning) with a modified 22G 5up cannula (Plastics One), which was then fed through the fascia layer over the shoulder and cemented next to the electrode implant site on the skull. In the same surgery, electrodes (drivable bundles of 10-25 µm diameter FeNiCr wires) were implanted dorsal to DLS (1 mm anterior to bregma, + or - 3.2 mm laterally, and 3.5 mm ventral to brain surface). Electrodes were advanced daily (40-80µm). We recorded 935 DLS neurons, 565 from 6 control rats (n's = 62, 82, 83, 88, 119, and 131 neurons) and 370 from 5 cocaine-exposed rats (n's = 19, 43, 81, 82, and 145 neurons).

Twelve-day self-administration protocol. After rats recovered from surgery, a twelve-day self-administration protocol was implemented using Med Associates Inc. operant behavioral boxes. During days 1-6, the cocaine group (n = 5) self-

administered 1mg/kg dosage of cocaine via lever press with a maximum of 30 infusions or 3-hour time limit. During days 7-12 a 0.5mg/kg dosage of cocaine was self-administered with a maximum of 60 infusions or a 3-hour time limit. Sessions began with illumination of house lights, extension of the lever, and illumination of a cue light above the lever. Active lever presses were paired with the cue light above the lever, which stayed on for the duration of the cocaine infusion (2.3 seconds per active infusion). Active lever presses for drug infusion could only take place 20 seconds after the previous active lever press. Lever presses during the inactive period resulted in no cue-light illumination or reward delivery. The active lever was extended for the duration of the session. Rats were taken out of the operant boxes after the maximum infusion or maximum time limit was reached as described above.

There were two control groups. One group received electrodes and catheters as did the cocaine group. This group (n = 2) followed the same protocol as the cocaine group except rats received sucrose pellets (Test Diet, 45 mg; 1-6d = 2 pellets; 7-12d = 1 pellet). All aspects remained the same as per the cocaine group regarding session start, active lever presses, cue light, and duration of cue light. The other control group (n = 4) only received electrodes and did not self-administer sucrose pellets. The latter group's data was published previously (Burton et al., 2014a).

Behavioral Analysis. Behavior during self-administration was evaluated by computing the average number of lever presses daily across rats in each group (cocaine and control) and computing the average number of lever presses across days 1-6 and days 7-12 across rats in each group. Behavior during performance of the task was evaluated by computing the percent choice of value conditions (short, long, big,

small) on free-choice trials as well as percent correct and reaction time (odor offset to odor port exit) on forced-choice trials for each value condition (short, long, big, small) for the first ten and last ten trials of each block of trials (total of four blocks per session). Percent choice was computed by determining how often the rat chose the more and less valued option on free-choice trials. Percent correct was computed by determining if the direction the animal chose was the same that was instructed by the forced-choice odor. Response bias for each session and each rat in each group was computed for free- and forced-choice trials by subtracting percent low-value choice from high-value choice (divided by the sum) for each session and by subtracting percent correct on low-value trials from percent correct on high-value trials (divided by the sum) for each session, and then averaging over the two.

Neural Analysis. Recorded waveforms were extracted from active channels during recording sessions and recorded to disk by an associated workstation with event time-stamps from the behavioral computer. Extracted single-units were then sorted in Offline Sorter using template matching (Plexon) and exported to NeuroExplorer to determine time-stamped events related to spike activity. All further analysis was done using Matlab. Analysis epochs were computed by taking the total number of spikes and dividing by time. That analysis epoch was taken from odor onset to odor port exit. The baseline epoch was one second before odor onset. Increasing- and decreasing-type neurons were categorized by whether or not they significantly increased or decreased firing compared to baseline, respectively ($p < 0.05$). A multifactor analysis of variance (ANOVA; $p < 0.05$) was performed for each increasing- and decreasing-type neuron to determine if activity was modulated

by stimulus (free- vs. forced-choice odors), response direction (contralateral vs. ipsilateral), and expected outcome (short, long, big, and small). Chi-square tests ($p < 0.05$) were performed to assess differences in the counts of neurons showing significant modulation across control and cocaine-exposed rats.

RESULTS

Self-administration

All rats were trained on the reward-guided decision-making task (Figure 1A) prior to implantation of electrodes in DLS (Figure 1G and H) and catheters for cocaine self-administration (see methods for more detail). During performance of the reward-guided decision-making task, on each trial, rats responded to one of two adjacent wells after sampling an odor at a central port (Figure 1A). Rats were trained to respond to three different odor cues: one odor that signaled reward in the right well (forced-choice), a second odor that signaled reward in the left well (forced-choice), and a third odor that signaled reward at either well (free-choice). Across blocks of trials in each recording session, we manipulated either the length of the delay preceding reward delivery (Figure 1A; Blocks 1-2; ~60 trial/block) or the size of the reward (Figure 1A; Blocks 3-4; ~60 trial/block).

Two weeks post-surgery, rats self-administered sucrose pellets or cocaine over the course of twelve days. During days 1-6 (1mg/kg cocaine or 2 sucrose pellets), the average number of active lever presses was 21.9 (±8.6; standard deviation) and 28.5 (±5.2; standard deviation) for cocaine and sucrose, respectively. During days 7-12 (0.5mg/kg cocaine or 1 sucrose pellet), the average number active lever presses were

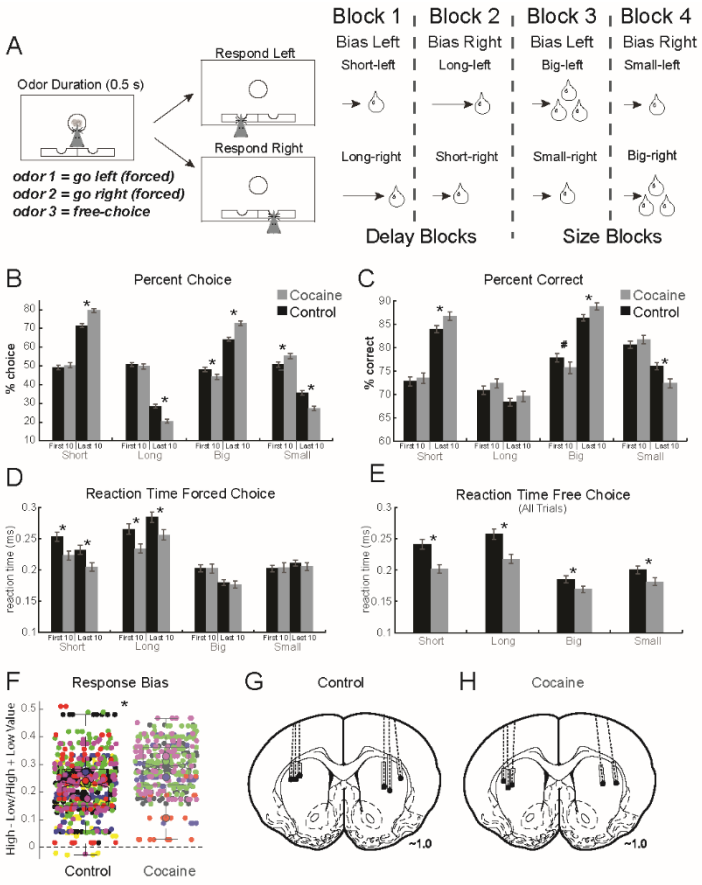

Figure 1: **A.** Task schematic, showing sequence of events in one trial (left panels) and the sequence of blocks in a session (right). Rats were required to nose-poke in the odor port for 0.5 s before the odor turned on for 0.5 s instructing them to respond to the adjacent fluid wells below where they would receive liquid sucrose reward after 500-7000 ms. For each recording session, one fluid well was arbitrarily designated as short (a short 500-ms delay before reward) and the other designated as long (a relatively long 1- to 7-s delay before reward) (Block 1). After the first block of trials (~60 trials), contingencies unexpectedly reversed (Block 2). With the transition to Block 3, the delays to reward were held constant across wells (500 ms), but the size of the reward was manipulated. The well designated as long during the previous block now offered 2–3 fluid boli, whereas the opposite well offered 1 bolus. The reward stipulations again reversed in Block 4. **B.** Percent choice on free-choice trials in each value manipulation over the first ten and last ten trials of each block averaged across animals and sessions (controls, black bars; cocaine, gray bars). **C.** Percent correct on forced-choice trials in the same manner as B. For these analyses, the sample was behavior associated with each unit to better represent changes in behavioral output as it relates to the firing rates of single neurons that contribute to the single-unit and population analysis described below. **D.** Reaction time (odor port exit minus odor offset) on forced-choice trials in the same manner as B and C. **E.** Reaction time (odor port exit minus odor offset) on all free-choice trials for each value manipulation. **F.** Response bias for each session and each rat in each group computed for free- and forced-choice trials by subtracting percent low-value choice from high-value choice (divided by the sum) for each session and by subtracting percent correct on low-value trials from percent correct on high-value trials (divided by the sum) for each session, and then averaging over the two. Each color represents a single rat; small dots represent each session and large dots represent the median overall all sessions within one rat. (*t test, $p < 0.05$; #, $p = 0.06$, error bars indicate SEM). **G-H.** Location of recording sites (Paxinos and Watson). Gray boxes mark the extent of the recording locations.

46.5 (±14.3) and 60 (±0) for cocaine and sucrose, respectively. After a month-long withdrawal period, rats were placed back in behavioral boxes interfaced with Plexon recording systems and recording commenced for approximately two months.

Cocaine biased behavior in the direction of high value rewards

After a month-long withdrawal period, activity in DLS was recorded during performance of the task. During these sessions, rats that had self-administered cocaine exhibited stronger response biases toward higher value reward. Control and cocaine groups tracked value across trials blocks choosing short-delay and large-reward more often than long-delay and small-reward on free-choice trials, respectively (Figure 1B). In an ANOVA with group (sucrose control, non-sucrose control, and cocaine), value (high and low), value manipulation (size or delay), and phase of learning within each block (early: first 10 trials; late: last 10 trials) as factors there was a significant main effect of value ($F(1,7456) = 3443$, $p < 0.05$) and no interaction between value and value manipulation ($F(2,7456) = 0.06$, $p = 0.94$). There was a significant interaction between group and value ($F(2,7456) = 60.8$, $p < 0.05$), with rats in the cocaine group choosing the high-value reward more often than controls in the last ten free-choice trials during both size (Figure 1B; $t(933) = 10.51$, $p < 0.01$) and delay (Figure 1B; $t(933) = 8.95$, $p < 0.01$) manipulations.

Consistent with the bias on free-choice trials, rats in the cocaine group were more strongly drawn to the high-value reward on forced-choice trials. In the ANOVA with percent correct as the dependent variable, there was a significant main effect of

value ($F(1,7456) = 376.49$, $p < 0.05$) and value manipulation ($F(1, 7456) = 258.49$, $p < 0.05$), as well as a significant interaction between group, value, and value manipulation ($F(2, 7456) = 807.4$, $p < 0.05$). During the last ten forced-choice trial types per block, rats that had previously self-administered cocaine were significantly more biased toward the side that produced better reward as evidenced by higher percent correct scores on short-delay (Figure 1C; $t(933) = 3.26$, $p < 0.01$) and large-reward forced-choice trials (Figure 1C, $t(933) = 3.10$, $p < 0.01$) and lower percent correct scores on small-reward forced-choice trials (Figure 1C; $t(933) = 4.00$, $p < 0.01$).

Exaggerated response bias observed during both free- and forced-choice did not emerge as a result of faster block switching as there were no differences between groups during the first ten trials of delay blocks (Figure 1B; Short: $t(933) = 0.99$, $p = 0.32$; Long: $t(933) = 0.99$, $p = 0.32$; Figure 1C; Short: $t(933) = 0.65$, $p = 0.51$; Long: $t(933) = 1.47$, $p = 0.14$). Further, cocaine-exposed rats were actually slower to reverse contingences early in size blocks, choosing small reward more often than large reward in the first ten free-choice trials (Figure 1B; $t(933) = 5.44$, $p < 0.01$) and trending towards diminished accuracy during the first ten large-reward forced-choice trials (Figure 1C; $t(933) = 1.85$, $p = 0.06$).

Overall, cocaine self-administration decreased reaction times (odor offset to nose poke exit); rats in the cocaine group responded to the odor significantly faster than controls on forced-choice trials in delay blocks and on free-choice trials during both delay and size blocks. In the ANOVA with reaction time as the dependent variable there was a significant main effect of group ($F(2,7456) = 1223.36$, $p < 0.05$)

and no interaction between group and value ($F(2,7456) = 0.11$, $p = 0.89$). There was also a significant group by value manipulation interaction ($F(2,7456) = 124.54$, $p < 0.05$) and a significant value by value manipulation interaction ($F(2,7456 = 11.3$, $p < 0.05$); cocaine rats were faster on forced choice trials during delay blocks during the first 10 trials (Figure 1D; Short: $t(933) = 3.13$, $p < 0.01$; Long: $t(933) = 2.93$, $p < 0.01$) and the last 10 trials (Figure 1D; Short: $t(933) = 2.96$, $p < 0.01$; Long: $t(933) = 2.58$, $p < 0.05$). Cocaine exposed rats did not exhibit different reaction times on forced-choice trials during performance of size blocks during the first 10 trials (Figure 1D; Big: $t(933) = 0.07$, $p = 0.95$; Small: $t(933) = 0.22$, $p = 0.830$) or the last 10 trials (Figure 1D; Big: $t(933) = 0.46$, $p = 0.65$; Small: $t(933) = 0.87$, $p = 0.38$). Finally, over all free-choice trial-types, cocaine exposed rats were faster relative to controls (Figure 1E; Short: $t(933) = 3.87$, $p < 0.01$; Long: $t(933) = 3.61$, $p < 0.1$; Big: $t(933) = 2.27$, $p < 0.05$; Small: $t(933) = 2.06$, $p < 0.05$).

Thus, overall, rats exposed to cocaine were biased toward higher value reward locations on both free- and forced-choice trials. This is further illustrated in Figure 1F, which plots a combined response bias measure for each session and each rat in each group. The response bias index was computed for the last ten free- and forced-choice trials by subtracting percent low-value choice from high-value choice (divided by the sum) for each session and by subtracting percent correct on low-value trials from percent correct on high-value trials (divided by the sum) for each session, and then averaging over the two. Small dots represent each session and large dots represent the average over all sessions within one rat. Across sessions the bias index was significantly larger in rats that had self-administered cocaine, consistent with the

analysis described above ($t(933) = 15.1$, $p < 0.01$). Only 1 rat (pale red) from the cocaine group fell below the median control response bias. Also, note that within the control group all rat averages fell between the lower and upper quartile, demonstrating that sucrose self-administration (sucrose controls, black and blue dots) did not impact response biases observed during task performance.

DLS firing reflects action-outcome contingencies divorced from the action selected after cocaine self-administration

In control and cocaine-exposed rats, 126 (22%) and 100 (27%) neurons increased and 262 (47%) and 124 (34%) decreased firing during odor sampling (odor onset to port exit) compared to baseline (1 second before odor onset; Wilcoxon; $p < 0.05$), respectively. The frequency of neurons that increased responding did not differ between groups (chi-square = 1.5; $p = 0.23$), however there were significantly fewer that decreased responding in rats that had self-administered cocaine (chi-square = 6.1; $p < 0.05$). To determine how the firing of these neurons were modulated during task performance, we performed an ANOVA ($p < 0.05$) with outcome (short, long, big, or small), stimulus-type (free or forced odor), and response direction (contralateral or ipsilateral) as factors on firing during the time between odor onset and odor port exit on trials in which reward was delivered (i.e., correct trials only). Initially we hypothesized that chronic cocaine use would amplify correlates related to stimulus-response processing by increasing the counts of neurons exhibiting significant main effects of stimulus or response direction, or interactions between stimulus and

response direction. Instead, we found an increase in neurons that exhibited an interaction between response direction and expected outcome (Figure 2A). The only significant group difference was that more neurons exhibited a significant interaction between response and outcome in cocaine compared to control rats (Figure 2A, chi square = 4.08, $p < 0.05$). This occurred only in neurons that increased firing during the odor epoch compared to baseline. Although neurons that decreased firing were selectively modulated by task parameters, the counts of neurons showing significant main or interaction effects did not differ between control and cocaine-exposed rats (Figure 2B).

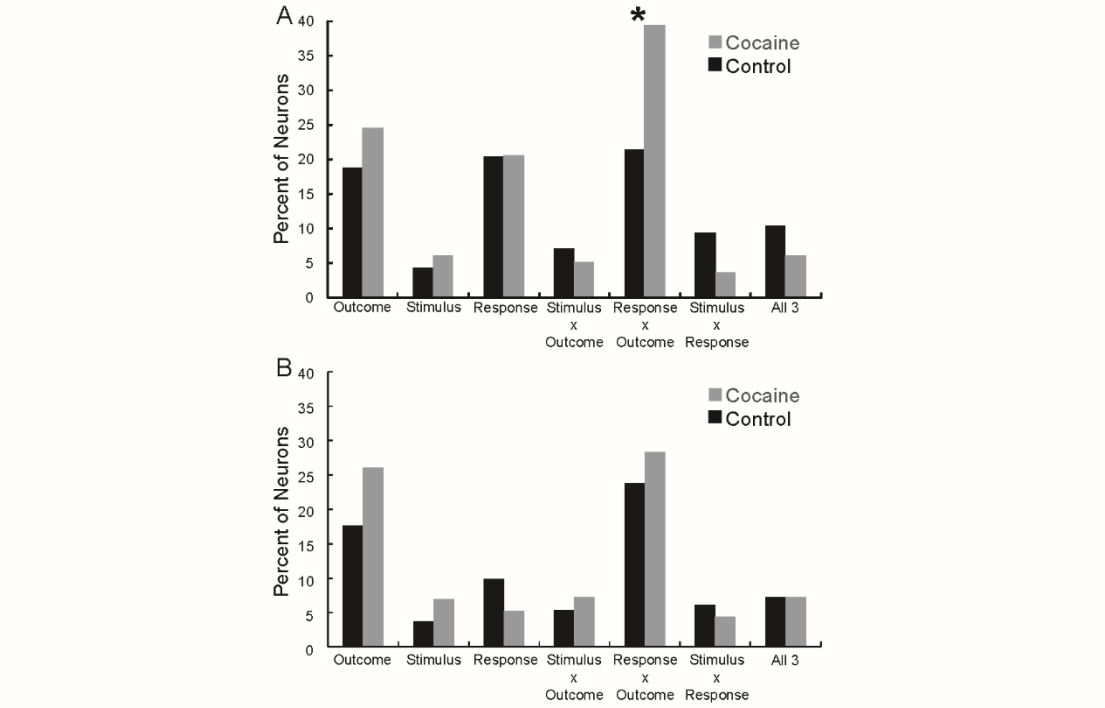

Figure 2: Percent of increasing- ((**A**) control, n = 126 (22%); cocaine, n = 100, (27%)) and decreasing-neurons (**B**) (control, n = 262 (47%); cocaine, n = 123(34%)) whose activity was significantly modulated by outcome, stimulus, response, and interactions between these three factors in the ANOVA. (*chi-square, $p < 0.05$).

As described in the introduction, in animals performing a two-choice reward paradigm, two different response-outcome correlates emerge, one that reflects the association between the action selected and the predicted outcome, and the other that reflects the contingency between reward and response direction independent of the action selected. An example of the former is illustrated by the firing of the single neuron plotted in Figure 3A. Neurons, such as this one, encode the relationship between the selected action (contralateral or ipsilateral) and the outcome predicted (short, long, big, or small). This particular neuron had a response field contralateral to the recording electrode (response field illustrated by dashed circles), firing strongly for actions made in the contralateral direction (top panels in Figure 3A). In addition, it was outcome selective, firing the most when the selected contralateral action would result in the delivery of reward after a short delay (top left panel in Figure 3A). Thus, this neuron conveyed information about the direction the rat was to select and the outcome that was predicted by that selection. We will refer to this as 'chosen-outcome' encoding.

Other neurons in DLS that exhibited a significant interaction between response and outcome did not encode the outcome predicted by the action selected, but instead the outcome predicted if the rat was to move in a particular direction independent from whether or not the animal actually chose to move in that direction. Take for example the firing of the neuron shown in Figure 3B: this neuron fired strongly in blocks of trials where the short delay is on the contralateral side, both when the rat will make a response toward the contralateral fluid well to obtain a reward after a short delay *and* when the rat will make a response to the ipsilateral

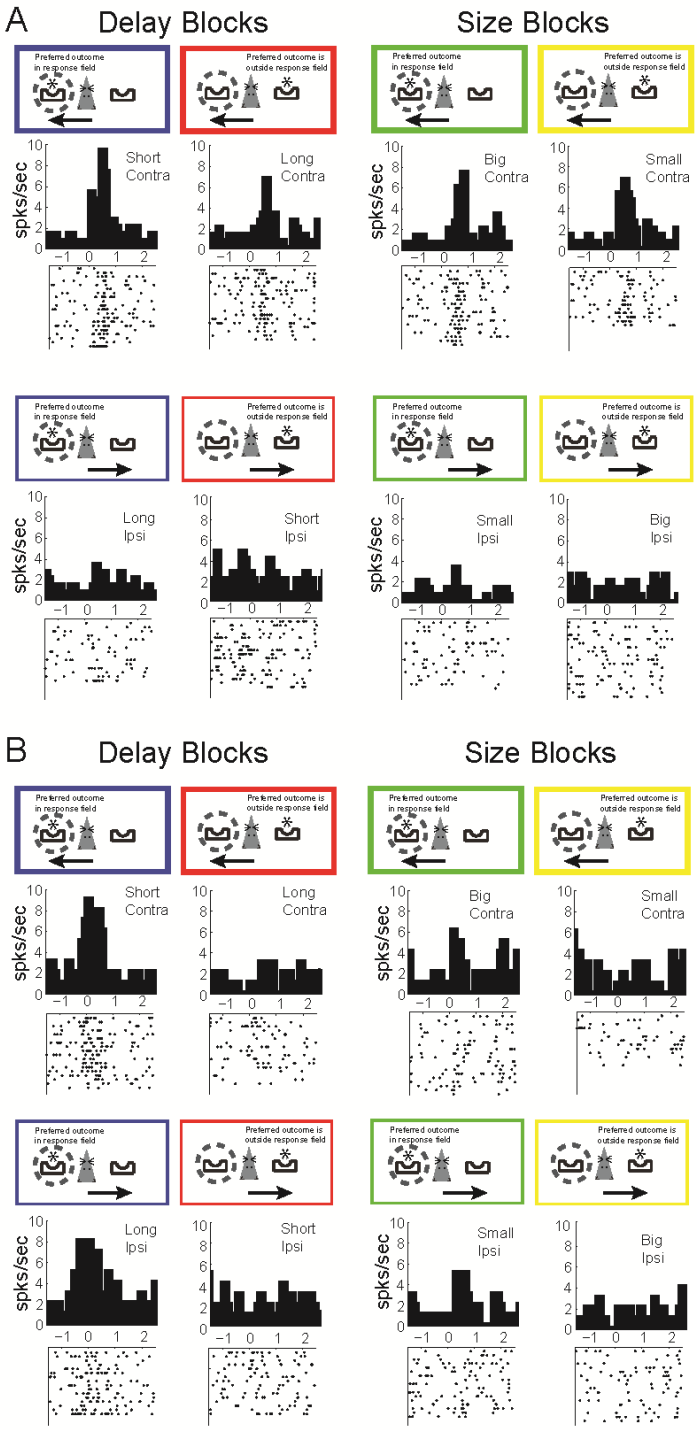

Figure 3: Single neuron examples of 'chosen-outcome'(**A**) and 'action-outcome' (**B**) encoding. Dashed circles around fluid wells in the colored boxes indicate the neuron's response field. Asterisks in the colored boxes refer to the location of the preferred outcome (inside or outside response field). Arrows within the colored boxes refer to the direction in which the animal made a response (into or away from response field). Thick and thin lines represent movements into and away from the response field. Activity is aligned to odor onset (0 s) and binned at 100 ms. Each tick mark equals 1 action potential.

fluid well to obtain reward after a long delay (blue panels, Figure 3B). Thus, activity of this neuron conveyed information that the short-delay (a high-value reward) is in the contralateral direction, independent of the action that was subsequently chosen. We will refer to this correlate as an 'action-outcome' signal. In the following paragraphs, we will show that action-outcome correlates, as defined here, are over-emphasized in the DLS after cocaine self-administration both at the population level and in the counts of single neurons.

The average population firing for control and cocaine-exposed rats is illustrated in Figure 4. In these plots, each neuron's preferred and non-preferred outcome and direction was determined by the neuronal response that elicited the most activity (spikes/second) during cue sampling (odor onset to odor port exit). In this figure, 'blue' reflects activity in the neuron's preferred block of trials (i.e., when the preferred outcome was in the response field as represented by the asterisk in the dashed circle), and 'red' reflects blocks of trials when the preferred outcome was outside of the response field (i.e. the non-preferred outcome). 'Green' and 'yellow' represent blocks of trials when the outcome of the same or opposite value (compared to blue and red, respectively) was in or outside the neuron's response field (e.g., if the preferred outcome was short, then the 'same' and 'opposite' value would be large and small, respectively). Individual trial-types are color-coded in Figure 3 as an example of how each neuron's response patterns fit into this color scheme.

Like the firing of the single-cell example shown in Figure 3A, population firing in the DLS of controls was highly selective. In controls, firing was significantly different during odor sampling for responses made into, but not away

from, the response field for the preferred outcome (Figure 4A and 4B; significance illustrated by SEM ribbons and running t-test; $p < 0.01$), reflecting the relationship between the selected action and the predicted outcome. This was not true after cocaine self-administration. Like controls, activity was also higher for responses made into the response field for the cell's preferred outcome (Figure 4C); however, activity was also significantly higher for behavioral responses made away from the neuron's response field in the same block of trials (Figure 4D). Thus, the population activity in cocaine exposed rats reflected the location of the preferred outcome in a particular context and did not reflect the outcome selected as in controls. That is, activity was high when the preferred outcome was in the cell's response field independent of the direction that the rat would eventually move, similar to the single-unit example in Figure 3B. Notably, selectivity in cocaine-exposed rats emerged prior to cue onset (black tick marks before zero) consistent with the idea that activity in DLS reflected the contingencies between actions and outcomes in a block of trials, which are known to the rat prior to cue onset.

The average population histogram suggests that there is an overabundance of action-outcome encoding neurons after cocaine self-administration. To determine if this was true, we performed a two-factor ANOVA with value manipulation (size or delay) and response-bias (contralateral response associated with high-value reward versus ipsilateral response associated with high-value reward, irrespective of value manipulation) as factors ($p < 0.05$). Neurons that show a main effect of response-bias without an interaction with value manipulation would reflect stronger firing whenever high-value outcomes (short delay or large reward) were in the response field (i.e.,

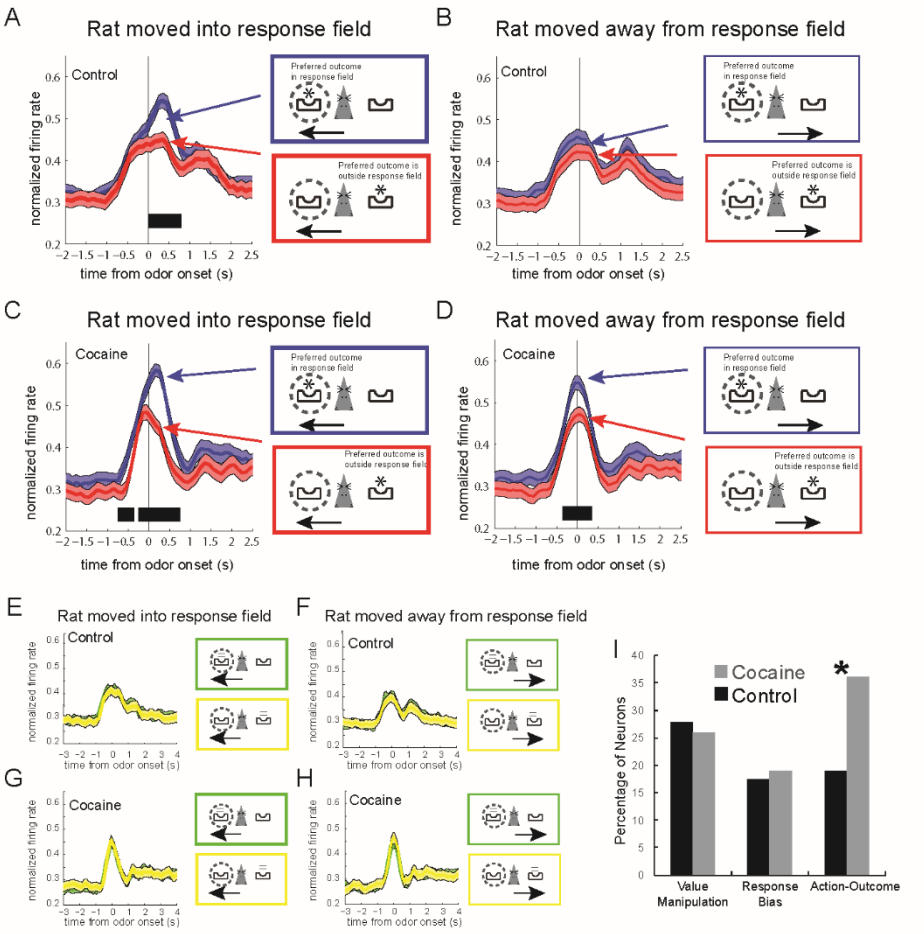

Figure 4: Action-outcome encoding is enhanced at the population level in cocaine-exposed rats. **A-B**. Population average for control rats (n = 126). A, trials in which the rat moved into the neuron's response field (preferred direction; thick lines). B, trials in which the rat moved away from the neurons response field (non-preferred direction; thin lines). Blue and red represent when the preferred and non-preferred outcome was in the response field, respectively. Dashed circles around fluid wells in the colored boxes indicate the neuron's response field. Asterisks in the colored boxes refer to the location of the preferred outcome (inside or outside response field). Activity is aligned to odor onset (0 s). Significance is illustrated by SEM ribbons and black tick marks (running t-test of two consecutive 100 ms bins; $p < 0.01$) under each histogram. **C-D.** Average firing over time for cocaine-exposed rats (n = 100) for response made into (C) and away (D) from the response field. Firing was high when the preferred outcome was in the response field (blue) regardless of whether the rat moved into (A) or away (B) from the response field. Neural activity is plotted in the same manner as Figure 4A-B. **E-H.** Histograms of neural population activity in DLS of control rats (E-F) and cocaine-exposed rats (G-H) for the same- and opposite-value outcomes (green and yellow colored lines and boxes, respectively) respective to Figure 4A-D. Neural activity is plotted in the same manner as Figure 4A-D. Histograms are normalized by dividing the max within each neuron. **I.** Proportion of neurons showing main effects of value manipulation (size or delay), response-bias (contralateral response associated with high-value reward, independent of value manipulation), or an interaction between these two factors (action-outcome) in ANOVA ($p < 0.05$). Cocaine-exposed rats had an increased number of neurons that exhibited action-outcome encoding compared to control rats (*chi-square, $p < 0.05$).

response-bias or action-value). Neurons that exhibit an interaction would reflect 'action-outcome' contingencies, not 'response-bias' or 'action-value', in that firing would only be higher for high-value outcomes within one value manipulation (size or delay).

The results of this analysis are shown in Figure 4I. The proportions of neurons that showed main effects of value manipulation and response bias exceeded chance levels in both groups (chi-squares > 40.97, ps < 0.05), but the frequency of main effects between cocaine-exposed and control rats did not significantly differ (value manipulation: cocaine = 26%; control = 28%; chi-square = 0.007, p = 0.94; response bias: cocaine = 19%; control = 17%; chi-square = 0.006, p = 0.94). However, as predicted by the population histograms (Figure 4A-H) and the single unit analysis described in Figure 3B, the counts of neurons that exhibited a significant interaction between value manipulation and response bias in cocaine animals (i.e., action-outcome) significantly outnumbered those in controls (Figure 4I; cocaine = 36%; control = 19%; chi-square = 4.11, p < 0.05).

Action-outcome signals emerged earlier and were not stimulus specific after cocaine self-administration

From the above analysis, it is clear that prior cocaine self-administration induces stronger action-outcome selectivity. To determine if this selectivity emerged earlier in a trial block in one group compared to the other we plotted trial-by-trial neural activity during odor sampling when the preferred outcome was in the response

field versus when the non-preferred outcome was in the response field. The analysis was performed separately for rewarded free- and forced-choice trials, and the trial at which activity was deemed selective was determined via a sliding t-test ($p < 0.01$; Figure 5, Gray ticks).

In controls, selectivity developed at trials 8 and 16 for free- and forced-choice trials, respectively. Notably, on free-choice trials, activity reflected the previous block's contingencies before reflecting the new action-outcome pairing during free-choice (Figure 5A (open ticks) in controls only. In cocaine-exposed rats, significant selectivity on free-choice trials emerged earlier than controls, arguably as early as trial 3, but convincingly by trial 7 (Figure 5B). Under forced-choice trials, significant differences were evident by trial 6 in cocaine-exposed rats (Figure 5D). Thus, overall, we conclude that action-outcome encoding emerged earlier in cocaine-exposed rats.

In the above analysis, preferred direction was defined separately for free- and forced- choice analysis. This was necessary because the firing of individual neurons might exhibit response-outcome selectivity for one but not the other trial-type (Stalnaker et al., 2010). From that analysis, it cannot be determined whether single neurons that exhibit action-outcome signals did so regardless of whether it was a free- or forced-choice trial-type. If these signals genuinely reflect action-outcome contingencies within a block of trials then one would predict that the activity of these neurons might reflect this relationship independent from the stimulus presented at the beginning of the trial. Certainly, the observation that selectivity emerged prior to stimulus onset (Figure 4) suggests that the effect is stimulus-independent.

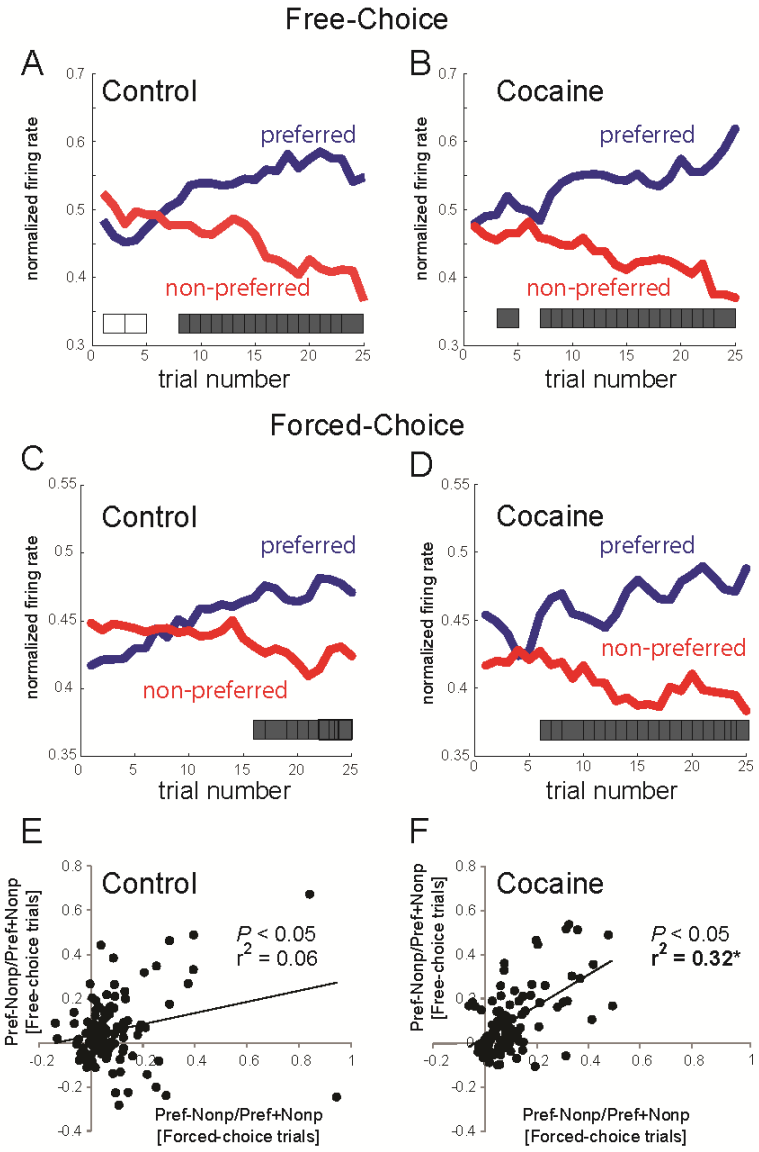

Figure 5: Neural selectivity emerges early in cocaine-exposed rats and was stimulus independent. **A-D.** Trial-by-trial normalized neural activity during odor sampling on correct free- and forced-choice trials for control rats (A and C) and cocaine-exposed rats (B and D) when the preferred outcome was in the response field vs. when the non-preferred outcome was in the response field. Tick marks underneath plots indicate two consecutive trials where there was a significant difference between preferred and non-preferred action-outcome contingencies (t-test, $p < 0.01$). **E-F.** Correlation between action-outcome selectivity (preferred minus non-preferred/preferred plus non-preferred) during free- and forced-choice trials. The neuron's preferred context was defined as the average spike activity of both free- and forced-choice trials that elicited the most activity during odor sampling. A positive correlation indicates that selectivity at the single neuron level was similar on free- and forced-choice trials.

To further address this issue, we performed an analysis where we defined the cell's preferred context based on the average of both free- and forced-choice trials. We then plotted the contingency (preferred minus non-preferred/preferred plus non-preferred) separately for forced- and free-choice trials (Figure 5E-F). Both groups exhibited a significant positive correlation (Figure 5E, control: $r2 = 0.06$, $p < 0.05$; Figure 5F, cocaine: $r2 = 0.32$, $p < 0.05$); however, the correspondence between free- and forced-choice trials was significantly stronger in the cocaine group compared to controls (Figure 5F, $z = 2.87$, $p < 0.05$). This demonstrates that selective firing observed within a particular block of trials occurs regardless of whether the cue was a free- or forced-choice odor.

DISCUSSION

Here we show that rats previously exposed to cocaine exhibit stronger behavioral response biases toward higher value reward and an overrepresentation of action-outcome signals at both the single neuron and population level. In cocaine-exposed rats, population firing in DLS did not reflect the outcome of the action that was about to occur ('chosen-outcome' encoding), but instead represented the location of the preferred outcome within the context of a particular block of trials, which we are referring to as 'action-outcome' encoding. Absent from this dataset was any evidence that S-R associations were enhanced after chronic cocaine self-administration during performance of our reward-guided decision-making task.

Our results support the hypothesis that cocaine self-administration alters response-outcome and model-based mechanisms, but demonstrates that correlates are

altered and not eliminated during decision-making. We show that population firing in DLS in cocaine-exposed rats failed to modify predictions based on the action selected, but instead showed an increase in neural activity reflecting the location of the preferred outcome regardless of subsequent instructions or movements in a particular block of trials. Further, action-outcome encoding in cocaine exposed rats was less stimulus-dependent, failing to take into account states within each trial that informed predictions about reward availability. In fact, action-outcome selectivity in cocaine exposed rats emerged *before* the onset of the odor. These altered neural correlates likely bias behavior successfully on free-choice trials, while diminishing performance on forced-choice trials. Overall, these results are consistent with a dysfunctional model-based system in that activity fails to represent the value of selected actions as governed by states in the task and over-represent contingencies between potential outcomes and actions within the context of the block, prior to and independent from the presented odor.

It might be argued that overrepresented action-outcome signals observed in cocaine exposed rats reflect some sort of exaggerated S-R encoding. This does not seem to be the case for several reasons. First, firing rate selectivity observed in DLS differed between size and delay blocks, even though S-R relationships were identical between them (i.e., smell this stimulus and make this response). Second, in the ANOVA performed on single neurons we did not observe any differences between control and cocaine-exposed animals in neurons showing selectivity for stimulus type or response direction, or an interaction between them (Figure 2). Further, after cocaine self-administration, we found that these correlates were actually less stimulus

bound, not more (Figure 5F), meaning that neurons fired selectively within a particular context regardless of whether the cue was a free- or forced-choice odor. Thus, this correlate does not appear to be modulated by stimuli and is outcome dependent, suggesting that it is not a form of S-R encoding.

It is also clear that other S-R correlates were not impacted by cocaine self-administration in that when we examined how many neurons were selective for stimulus type or response direction (or an interaction between them), selectivity was similar between cocaine exposed and control rats. This is true even though forced-choice contingencies (i.e., odor 1 = left, odor 2 = right) never changed over several months of training and recording. The fact that cocaine exposure did not impact stimulus and response processing was somewhat of a surprise to us because in a previous study we showed that lesions to nucleus accumbens (NAc) increased the counts of neurons that were selective for odors and responses in DLS, and enhanced the overall strength of the signal at the population level (Burton et al., 2014a). Considering that NAc is one of the first areas to be impacted by drug use, (Koob and Volkow, 2010; Steinberg et al., 2014; Keiflin and Janak, 2015; Lüscher, 2016), we expected a similar result after cocaine self-administration.

The lack of altered S-R correlates in our study likely reflects the complex nature of our two-choice paradigm where there is no simple mapping between stimuli, responses, and outcomes, rather than being due to differences in doses or withdrawal times. Not only does this paradigm require constant tracking of reward across two responses, but also an awareness of what rewards are available on each trial as conveyed by odor identity. In this situation, behavior is governed by multiple

action-outcome contingencies, which is known to encourage goal-directed behavior (Colwill and Rescorla, 1985; Dickinson et al., 2000; Colwill and Triola, 2002; Holland, 2004; Kosaki and Dickinson, 2010). Even though we did not classically devalue outcomes in our task and cannot definitively prove that rats are using representations of action-outcome contingencies, others have shown that under similar circumstances exposure to drugs of abuse does indeed leave goal-directed mechanisms intact and in some cases, enhanced (Phillips and Vugler, 2011; Son et al., 2011; Halbout et al., 2016). In particular, Halbout and colleagues showed that when rats are trained on two different action-outcome contingencies, rats exposed to cocaine stop responding to the one that was devalued and that after contingency degradation rats exposed to cocaine actually alter behavior more quickly than controls. Our findings are consistent with these results in that we do not see enhanced S-R encoding, do see increased action-outcome correlates, and rats exposed to cocaine make faster adjustments in behavior when contingencies change.

We conclude that prior cocaine-exposure increases action-outcome processing without impacting chosen-outcome and S-R encoding in our decision-making task. Here, we define action-outcome correlates as the relationship between actions and outcomes independent from that actual action that will be taken. In this way, it could be viewed that cocaine self-administration increases outcome encoding that is divorced from the action that will be selected. This correlate is more of a reflection of the contingencies available during decision-making as opposed to a representation of what will ultimately be selected. Such correlates can only be examined in the context of neural recording in a paradigm with at least two choices. Notably, representations

of response- or action-outcome associations as more classically defined in the learning theory literature (i.e., selected action-outcome or chosen-outcome) were not altered by cocaine self-administration. The increase in 'action-outcome' correlates that we describe here are more similar to 'response-bias' or 'action-value' signals described in the caudate of primates (Lauwereyns et al., 2002; Lau and Glimcher, 2007, 2008; Nakamura et al., 2012), which are thought to represent the association between actions and outcomes to push the motor system toward decisions that lead to high-value outcomes independent from instructed/selected movement. As in these reports, this signal emerged prior to instructional cues reflecting the context in which rewards were distributed prior to the decision period. Here, we show that similar signals, albeit outcome-dependent (i.e., specific to size or delay), are amplified after chronic cocaine use at the expense of correlates that inform behavior via computations of predicted outcomes based on upcoming decisions.

Chapter 4: Ventral striatum lesions enhance stimulus and response encoding in dorsal striatum

Published Manuscript

Biological Psychiatry. 2014 Jan 15;75(2):132-9. doi:10.1016/j.biopsych.2013.05.023.

Epub 2013 Jun 20.

Amanda C. Burton*[1,2] , Gregory B. Bissonette *[1], Nina T Lichtenberg[1] , Vadim Kashtelyan[1] and Matthew R. Roesch [1,2]

1. Department of Psychology
2. Program in Neuroscience and Cognitive Science
1147 Biology-Psychology Building
University of Maryland
College Park, MD 20742

*These authors contributed equally to this paper.

ABSTRACT

The development of addiction is thought to reflect a transition from goal-directed to stimulus-response driven behavior, functions attributed to ventral (VS) and dorsal striatum (DS), respectively. In line with this theory, neuroadaptations that occur during prolonged drug use progress from VS to DS. Here we ask if VS dysfunction alone, independent of drug use, can affect neural selectivity in DS. To address this issue, we recorded from single neurons in dorsal lateral striatum (DLS) while rats performed an odor-guided choice task for differently valued rewards in rats with and without unilateral VS lesions, specifically in the nucleus accumbens core (NAc). In a separate group of animals, we used bilateral NAc lesions to determine if NAc was critical for performance on this task. We describe data showing that unilateral lesions of NAc enhance neural representations in DLS during performance of a task that is dependent on NAc. Furthermore, we show that NAc is critical for reward-guided decision-making initially, but that rats regain function after several days. These results suggest that loss of NAc function, independent of chronic drug use, can trigger stronger encoding in DLS in a reward-guided decision-making task.

INTRODUCTION

Ventral (VS) and dorsal striatum (DS) perform critical roles in reward-guided decision-making and reinforcement learning, but it is still unclear how they interact. Together, with midbrain dopamine neurons, they form a circuit commonly referred to as the actor-critic model (Barto, 1995; Houk et al., 1995; Sutton and Barto, 1998; Haber et al., 2000; Joel et al., 2002; O'Doherty et al., 2004; Redish, 2004b; Ikemoto, 2007; Niv and Schoenbaum, 2008; Takahashi et al., 2009; Van der Meer and Redish, 2011; Padoa-Schioppa, 2011). In this model, VS and dopamine neurons function to generate reward predictions and prediction errors, which modify action policies in DS so that desired outcomes can be obtained. This circuit is thought to be critical for drug-seeking and is affected by chronic drug use (Everitt and Robbins, 2005, 2013; Hyman et al., 2006; Belin and Everitt, 2008; Koob and Volkow, 2010).

Many behaviors, including drug-seeking, are initially goal-directed, but eventually become stimulus driven or habitual with repetition. The transition away from goal-directed behavior toward stimulus driven habits is thought to depend upon a switch in control from VS to DS (Everitt and Robbins, 2005; Vanderschuren et al., 2005), which is amplified by drugs of abuse (Everitt et al., 2001; Everitt and Robbins, 2005). Since many structural and functional alterations occur with extended drug use it is still unclear what might initiate this change. Importantly, the NAc region of VS appears to be one of the earliest brain regions to be affected by administration of drugs of abuse, suggesting that its disruption might be enough to initiate changes in

downstream areas critical for stimulus driven behaviors such as dorsal lateral striatum (DLS).

Indeed, above we show that both NAc and DLS are both altered after cocaine self-administration. However, we found that stimulus-response correlates were unaltered in DLS, whereas correlates related to response-outcome encoding were enhanced and divorced from actions. That is, after cocaine exposure, correlates in DLS better reflected the contingencies available during decision-making as opposed to a representation of what would ultimately be selected. Notably, these correlates were amplified after chronic cocaine self-administration at the expense of correlates that inform behavior via computations of predicted outcomes based on upcoming decisions. This lack of processing of direction and value in DLS might have resulted from impaired encoding in NAc, where we observed diminished value and directional signaling.

Here, we ask if loss of NAc function alone, independent of drug use, might increase encoding in DLS. We found that stimulus and response encoding in DLS was enhanced after NAc lesions during performance of a task that was dependent on NAc. After several days of post -surgery training, lesioned rats were able to make accurate reward-guided decisions, suggesting that enhanced encoding in DLS might compensate for loss of NAc function. These results demonstrate that disruption of decision-making with lesions to NAc is enough to amplify S-R signals in DLS. While we did find disruptions in encoding in NAc, specifically diminished directional and value signaling, we found enhanced R-O associations in DLS after cocaine exposure.

This suggests that exposure to drugs of abuse is not as simple as a disruption in communication between the NAc and DLS.

MATERIALS AND METHODS

Subjects. Twenty-six male Long-Evans rats were obtained at 175–200g from Charles River Labs. Rats were tested at the University of Maryland, College Park in accordance with UM and NIH guidelines.

Surgical procedures. All surgical procedures were performed after training on the task described below. Ten rats had a drivable bundle of 10–25 μm diameter FeNiCr wires chronically implanted in the left or right hemisphere dorsal to DLS (n = 10; 1 mm anterior to bregma, + or − 3.2 mm laterally, and 3.5 mm ventral to the brain surface) (Singh et al., 2010; McDannald et al., 2011; Bryden et al., 2012). NAc lesions were made with a 2 μl Hamilton syringe, beveled edge facing the posterior direction, using 0.11M quinolinic acid, pH 7.4 in Dulbecco's PBS (Sigma). Quinolinic acid (0.3 μl) was delivered at 0.15 μl/min at coordinates: AP +1.9 ML ±1.9 DV −7.3. The remaining 4 rats served as controls, which received sham surgeries during which the Hamilton syringe loaded with saline was lowered to the same coordinates. In addition to rats that received electrodes, another group of rats only received bilateral sham (n = 6) or NAc lesions (n = 8) to characterize behavior. Brains were removed and processed for histology using standard techniques at the end of the experiment.

Odor-guided decision-making task. Before surgery, all rats were trained on the odor-guided delay/size choice task. On each trial, nose poke into the odor port after

house light illumination resulted in delivery of an odor cue to a hemicylinder located behind this opening (Bryden et al., 2011b; Roesch and Bryden, 2011). One of three different odors (2-Octanol, Pentyl Acetate, or Carvone) was delivered to the port on each trial. One odor instructed the rat to go to the left to receive reward, a second odor instructed the rat to go to the right to receive reward, and a third odor indicated that the rat could obtain reward at either well. Odors were presented in a pseudorandom sequence such that the free-choice odor was presented on 7/20 trials and the left/right odors were presented in equal proportions.

During the first day of training rats were first taught to simply nose poke into the odor port, then respond to the well for reward. On the second day, the free-choice odor was introduced and rats were free to respond to either well for reward. On each subsequent day, the number of forced-choice odors increased by 2 for each block of 20 trials. During this time, we introduced blocks in which we manipulated the reward size and the length of the delay preceding reward. Once the rats were able to maintain accurate responding (> 65%) on forced-choice trials through these manipulations, surgery was performed.

During recording, one well was randomly designated as short (500 ms) and the other long (1–7s) at the start of the session (Figure 1A: Block 1). In the second block of trials, these contingencies were switched (Figure 1A: Block 2). The length of the delay under long conditions abided by the following algorithm: the side designated as long started off as 1s and increased by 1s every time that side was chosen on a free-choice odor (up to a maximum of 7s). If the rat chose the side designated as long less than 8 out of the previous 10 free choice trials, the delay was

reduced by 1s for each trial to a minimum of 3s. The reward delay for long forced-choice trials was yoked to the delay in free-choice trials during these blocks. In later blocks, we held the delay preceding reward delivery constant (500 ms) while manipulating the size of the expected reward (Figure 1A: Blocks 3 and 4). The reward was a 0.05 ml bolus of 10% sucrose solution. For big reward, an additional bolus was delivered 500 ms after the first bolus. Essentially there were four basic trial types (short, long, big, and small) by two directions (left and right) by two stimulus types (free- and forced-choice odor).

For behavior after bilateral lesions, only one manipulation varied each day. On each day, one well was randomly designated as high value (i.e. short delay or large reward depending on the day). The location of the high value outcomes switched every 60 correct trials. There was a total of 3 blocks each day. Delay manipulations occurred on days 1, 3, 5 and 7. Size manipulations occurred on days 2, 4, 6, and 8. All other contingencies were the same as during recording.

Single-unit recording. Procedures were the same as described previously. Electrodes were advanced daily (40–80μm). Neural activity was recorded using four identical Plexon systems interfaced with odor discrimination chambers. Waveforms (>2.5:1 signal-to-noise) were extracted from active channels and recorded.

Neural analysis. Analysis epochs were computed by taking the total number of spikes and dividing by time. That analysis epoch was taken from 100 ms after odor onset to completion of the behavioral response. The activity of neurons during this epoch for which we examined differences between trial types at the single cell level only violated normality (Jarque-Bera; $p < 0.05$) in 6% of the 809 recorded neurons,

which is fewer than expected from chance alone (chi-square; $p = 0.09$). A multi-factor ANOVA ($p < 0.05$) was performed for each neuron to determine if activity was modulated by odor type (free versus forced), response direction (left versus right), and expected outcome (short, long, big and small).

Neurons were also characterized by comparing firing rate during baseline to firing rate during the analysis epoch, averaged over all trial types (t test, $p < 0.05$). Baseline was the average firing rate taken for 1s, starting 2s before odor onset. Chi-squares were performed to assess differences in the counts of neurons showing significant modulation across lesion and controls.

Behavioral analysis. Behavior during performance of the task was evaluated by computing percent choice of high and low value outcomes on free-choice trials, and percent correct and reaction time (odor offset to odor port exit) on forced-choice trials. Multi-factor ANOVAs were performed on these behavior measures to assess differences between lesions and controls. Factors included value (high versus low), manipulation (delay vs size), lesion (sham vs lesion), day (1–4) and block (1–3).

RESULTS

Rats were trained on a reward-guided decision-making task previously used to characterize encoding in several areas (Figure 1A) (Stalnaker et al., 2010; Roesch and Bryden, 2011). Rats learned to nosepoke in a central odor port, wait for delivery of an odor (500 ms) and then respond to one of two fluid wells located to the left or right of the odor port. Odors signaled forced- and free-choice trials. On forced-choice trials, rats had to respond to the fluid well signaled by the odor (left or right) to receive

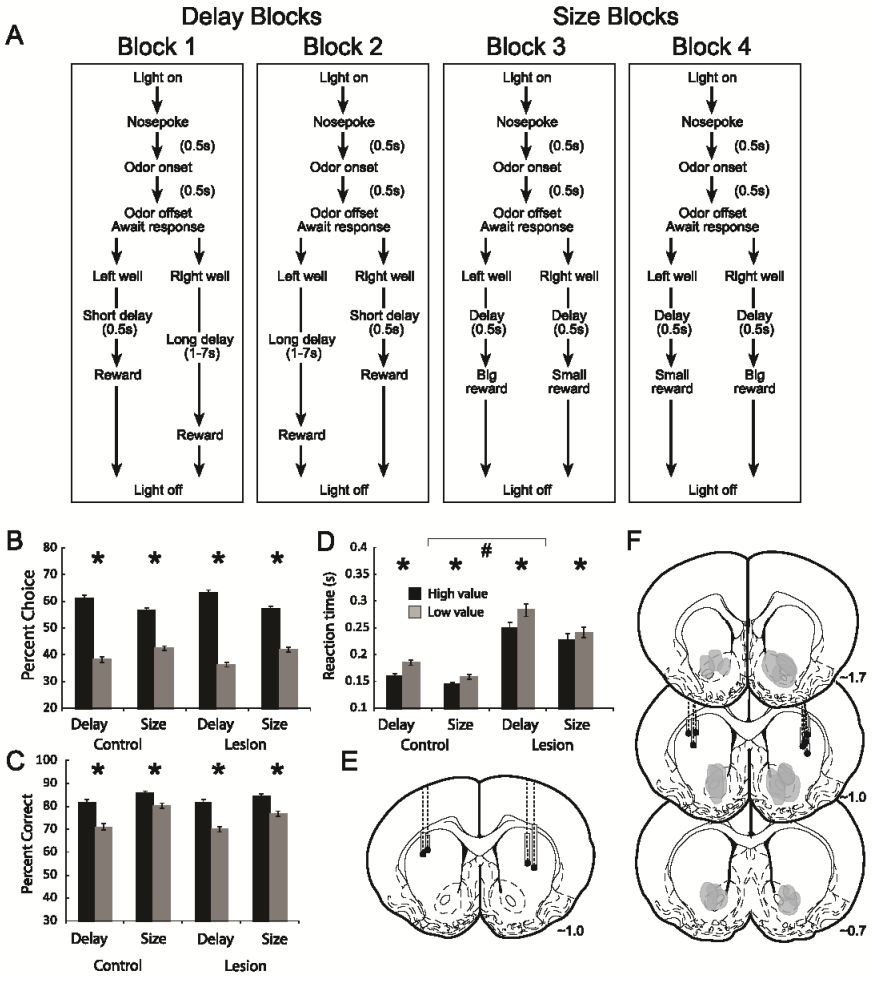

Figure 1. Task, behavior and recording/lesion locations. **A.** An example of the sequence of events in each trial block. For each recording session, one fluid well was arbitrarily designated as short (500ms delay before reward) and the other designated as long (1-7s delay before reward) (Block 1). After the first block of trials (~60 trials), contingencies unexpectedly reversed (Block 2). With the transition to block 3, the delays to reward were held constant across wells (500ms), but the size of the reward was manipulated. The well designated as 'long' during the previous block now offered 2-3 fluid boli whereas the opposite well offered one bolus. The reward stipulations again reversed in block 4. Free-choice odors signal that either well could be selected for reward, whereas forced-choice odors signaled that reward would only be delivered in the well that the rat was instructed to go to. **B**. The impact of delay length and reward size manipulations on choice behavior during free-choice trials. Percent choice is calculated by taking the number of choices made and divided by the total number of well entries on free-choice trials, multiplied by 100. **C.** Impact of value on forced-choice trials for short vs. long delay and big vs. small reward. **D**. Reaction times (odor offset to nose unpoke from odor port) on forced-choice trials comparing short vs. long delay trials and big vs. small reward trials. High value = short and large. Low value = long and small. **E-F**. Location of recording sites and unilateral lesions based on histology for sham (E) and lesioned rats (F). Recordings and lesions were performed in the same hemisphere (3 lefts; 4 rights). Filled gray boxes mark the locations of electrodes based on histology and initial recording site. Black dot marks the bottom of the recording tract. Transparent gray areas mark lesions for each animal. Shown are representative slices at 1.7, 1.0 and 0.7 anterior to bregma taken from Paxinos and Watson (1997). Asterisks indicate planned comparisons revealing statistically significant differences (t test, $p<0.05$). # indicates a main effect of lesion in the ANOVA ($p < 0.05$). Error bars indicate standard error of the mean (SEM).

reward. On free-choice trials rats were free to select either well. Over the course of four trial blocks we independently manipulated the length of the delay to (0.5 s or 1–7 s) or the size of (1 vs 2–3 boli) reward, making one fluid well better than the other. Essentially there were four predicted outcomes (short, long, big and small) by two directions (left and right) by two trial-types (free and forced odors).

After training on this task rats were split into two groups: in both groups, we implanted recording electrodes in DLS, in the lesion group (n = 7) rats received unilateral NAc lesions in the same hemisphere as the DLS electrode. Control rats received sham lesions (n = 4; see methods for further detail). Unilateral lesions were chosen to examine the impact of NAc lesions on DLS neural selectivity, with a minimal impact on behavior. Lesions and electrode positions are illustrated in Figure 1E–F.

Both control and lesion rats perceived differently delayed and sized rewards as having different values across all four trial blocks. There was a main effect of value with no interaction with lesion (ANOVA; $p < 0.05$). On free-choice trials, rats chose the well associated with large reward and short delay significantly more often than the well associated with small reward and long delay, respectively (Figure 1B; t test; p's < 0.05). On forced-choice trials, rats were more accurate and faster on large reward and short delay trials, as compared to their respective counterparts, small reward and long delay (Figure 1C–D; t test; p's < 0.05). Thus, performance on free- and forced-choice trials was modulated by the predicted outcomes in both size and delay trial blocks for both controls and lesions.

The only significant difference between the lesion and sham groups was that rats with NAc lesions were significantly slower to decide when to leave the odor port (reaction time = port exit minus odor offset) compared to controls (Figure 1D; control versus lesion; ANOVA; interaction effect; $p < 0.05$). Note, that this does not reflect a gross motor deficit because rats with lesions moved from the odor port to the fluid well (movement time) significantly faster than controls (t test; $p < 0.05$), suggesting that decisions took longer to process in rats with NAc lesions, but the ability to move or act on that decision was unimpaired.

Stimulus and response encoding in DS was enhanced after VS lesions

We recorded activity from 457 and 352 DLS neurons from lesion and control rats, respectively (Figure 1E and F). As described previously, neural activity in DLS was highly associative, being modulated by all aspects of the task: trial-type (free vs forced odors), response direction (left vs right) and expected outcome (short, long, big and small) (Stalnaker et al., 2010). This is illustrated by the activity of the single cell example in Figure 2A during performance of size blocks. This neuron fired the strongest for forced-choice odors that predicted large reward and was directionally tuned in that firing was stronger for movement in the direction contralateral to the recording site (i.e. right).

The average firing over all recorded DLS neurons from lesions and controls are shown in Figure 2B–C. To make this plot, we sorted trials based on preferred trial-type according to which condition produced the maximal response. The remaining trials were categorized relative to the preferred trial type depending on

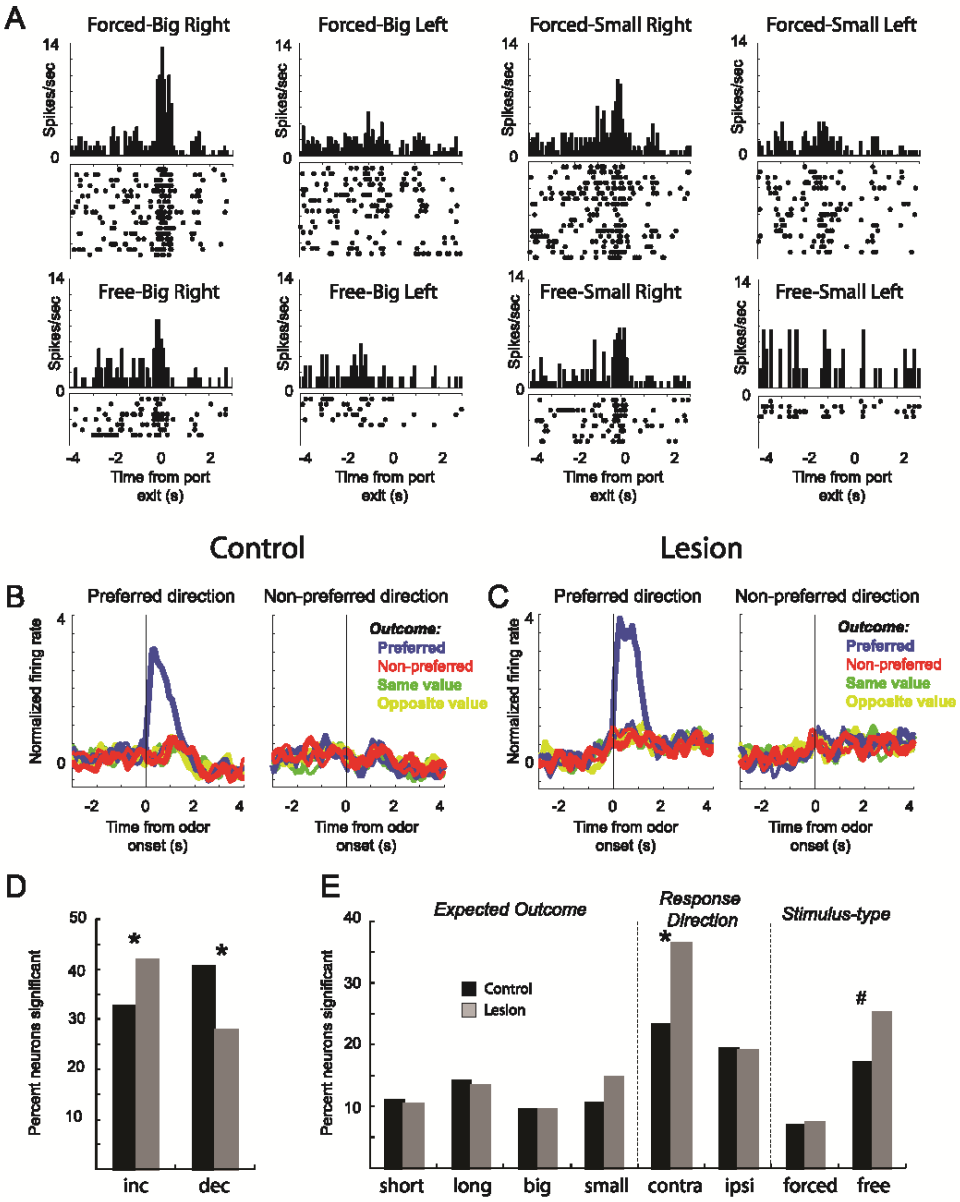

Figure 2. Stimulus and response encoding in DS was enhanced after VS lesions. A. Single cell example showing activity during size blocks during forced (top row) and free (bottom row) choice performance. **B-C**. Average firing rate over all 352 and 457 neurons for controls (b) and lesions (c) for free- or forced-choices trials, depending on which elicited the strongest response. All trials are referenced to the trial-type that elicited the maximal firing during the time starting 100 ms after odor onset to entry into the fluid well. See text for more detail. Data were normalized by subtracting the mean and dividing by the standard deviation (z score). Blue = preferred outcome; red = non-preferred outcome; green = same value and yellow = opposite value; thin = first 5 trials; thick = last 5 trials. **D**. Percent of neurons that showed significant increases or decreases, averaged across trial-type, during the analysis epoch compared to baseline (ttest; $p < 0.05$). **E**. Height of each bar indicates the percent of neurons that showed a main effect or interaction effect of outcome (short, long, big and small), direction (contra and ipsilateral) and odor-type (free and forced choice odors). * = $p < 0.05$; # = 0.11; chi-square.

whether the response and outcome were in the same or opposite direction, and were the same or opposite value for both manipulations (delay and size). For example, if a neuron fired maximally for big forced-choice trials to the right, then big reward and right became the preferred outcome and direction, small and left became non-preferred outcome and direction, and short and long became outcomes of the same and opposite value, respectively.

As described previously and consistent with the single cell example, population activity was highly selective, firing for a specific combination of odor, outcome and direction (Stalnaker et al., 2010). Although activity was strongest for the preferred outcome and direction (by definition), there was little modulation by other trial types (Figure 2B and C; blue versus other colors). This was true for both experimental and control groups. The only noticeable difference between population firing in sham and lesion rats is that cue-evoked activity appeared to be enhanced, and overall stronger, in rats with NAc lesions (Figure 2B versus 2C). Indeed, the difference between the trial type that elicited the strongest firing and the trial-type designated as its opposite value and direction was significantly stronger in lesions compared to controls (t test; 2.9 versus 2.2, $p < 0.05$). Thus, at the population level, neural selectivity in DLS was stronger in rats with NAc lesions.

Consistent with a general increase in firing observed after lesions, the count of neurons that significantly increased and decreased firing during the decision period were significantly different between the two experimental groups (Figure 2D). For this analysis, we compared average firing over all trial-types during the decision period (100 ms after odor onset to fluid well entry) to baseline (1 s prior to trial start; t

test; $p < 0.05$). For controls, the number of neurons that increased and decreased firing were 116 (33%) and 144 (40%), respectively. This proportion was flipped for lesions, with 193 (42%) and 129 (28%) showing increased and decreased firing, respectively (Figure 2D; chi-square; $p < 0.05$). The total number of responsive cells (increasing plus decreasing) were not significantly different between the two groups (control = 260(73%); lesion = 322(70%); chi-square; $p = 0.7$). Thus, the overall increase in firing observed at the population level reflects an altered frequency of increasing- and decreasing-type neurons in lesions compared to control rats.

To determine if selectivity at the single neuron level was different between lesions and controls we characterized firing by performing a ANOVA with stimulus (free- versus forced odors), response direction (contra vs ipsi), and predicted outcome (short, long, big, and small) as factors during the period starting 100 ms after odor onset and ending upon fluid well entry.

The results of this analysis for the entire population of responsive neurons are illustrated in Figure 2E. The height of each bar represents the percent of neurons that showed a significant ($p < 0.05$) main or interaction effect of outcome (short, long, big, and small), response direction (contra or ipsilateral to the recording site), and/or stimulus-type (free- and forced-choice odor). Each factor was then broken down by which trial-type produced the maximal firing.

As in our previous data set, an equal number of neurons from control rats fired maximally for each of the four predicted outcomes (Figure 2E; black bars; short, long, big and small) (Stalnaker et al., 2010). This effect was also present in rats with NAc lesions (Figure 2E; gray bars). The proportion of neurons selective for the four

outcomes did not significantly differ between the two groups (Figure 2E; short, long, big and small; chi-square; p's > 0.05).

Activity in DLS was highly directional, firing more strongly for one direction over another, as illustrated by the single cell example and the population. Remarkably, this directional response selectivity was enhanced in rats with NAc lesions, specifically for movements made in the contralateral direction relative to the recording site. The counts of neurons selective for contralateral movement in lesioned rats significantly outnumbered those selective for contralateral movement in controls (Figure 2E; response direction; contralateral; black vs gray; chi-square; $p < 0.05$).

We have also shown that activity in DLS is stimulus selective, with the majority of neurons showing maximal firing for free-choice odors (Stalnaker et al., 2010). Consistent with those findings, counts of neurons that differentiated between free and forced odors were more than expected by chance alone (Figure 2E; black bars; stimulus-type; chi-square; $p < 0.05$) and the frequency of neurons firing more strongly for free-choice odors were in the majority (Figure 2E; black bars; stimulus-type; chi-square; $p < 0.05$). Like directional response tuning, stimulus selectivity was enhanced in rats with NAc lesions, however, the elevated frequency of neurons showing increased firing in lesioned rats on free-choice trials only approached significance when examining the population as a whole (Figure 2E; #: chi-square, $p = 0.11$).

Next, we asked if enhanced stimulus and response encoding was consistent across neurons that showed general increases and decreases in firing during the decision period. Remember that the ratio of increasing and decreasing neurons flipped

after lesions, in that the number of increasing-type neurons were in the significant majority in the lesion, but not the control group (Figure 2D). The population analysis performed earlier on all cells is broken down by increasing- and decreasing-type in Figure 3. Remarkably, response direction and stimulus encoding enhancement observed across the entire population was a result of specific changes in increasing- and decreasing-type neurons, respectively. For increasing-type cells the frequency of neurons that showed elevated firing for contralateral movement outnumbered those observed in controls (Figure 3E; Response Direction; black versus gray; chi-square; $p < 0.05$). For decreasing cells (Figure 3F), there was a significant increase (chi-square; $p < 0.05$) in the number cells that were modulated significantly by either free- or forced-choice odors (ANOVA; $p < 0.05$), with a trend toward a disproportionate increase on free-choice trials (Figure 3F; #: chi-square; $p = 0.057$). We conclude that NAc lesions induced increased firing during task performance and enhanced selectivity for stimuli and responses at the single unit and population level.

Reward-guided decision-making is disrupted after NAc lesions

The results thus far suggest that NAc is not critical for outcome encoding in DLS and that associations related to direction and stimuli were enhanced after NAc lesions. Although deficits in behavior have been observed after NAc disruption, these results raise the important question of whether NAc is critical for performance on our task (Brown and Bowman, 1995; Cardinal et al., 2001; Giertler et al., 2003, 2004; Acheson et al., 2006; Floresco et al., 2008). If NAc is not critical, then we would not expect to see loss of outcome encoding in DLS in the first place.

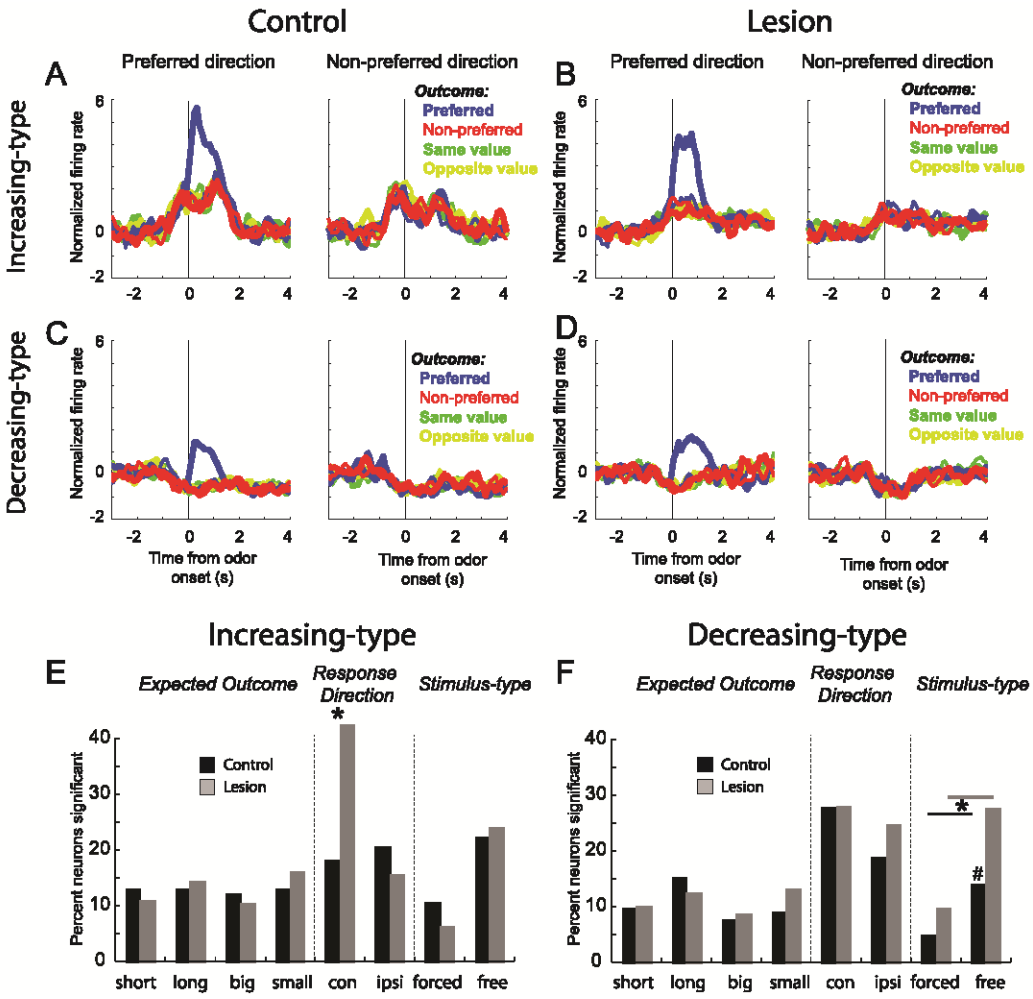

Figure 3. Stimulus and response encoding in decreasing- and increasing-type DS neurons was enhanced after VS lesions. A-D. Average firing rate over all increasing (A-B) and decreasing (C-D) type neurons for controls (A and C) and lesions (B and D) for free- or forced-choice trials, depending on which elicited the strongest response. All trials are referenced to the trial-type that elicited the maximal firing during the time starting 100 ms after odor onset to entry into the fluid well. See text for more detail. Data were normalized by subtracting the mean and dividing by the standard deviation (z score). Blue = preferred outcome; red = non-preferred outcome; green = same value and yellow = opposite value; thin = first 5 trials; thick = last 5 trials. **E-F**. Height of each bar indicates the percent of neurons that showed a main or interaction effect of outcome (short, long, big and small), direction (contra and ipsilateral) and odor-type (free and forced choice odors). * $p < 0.05$; # = 0.057; chi-square.

To address this issue, we trained a second group of animals in the same manner as the first, but instead of implanting electrodes we made bilateral lesions to NAc (Figure 4). We tested them for eight days, alternating delay and size manipulations. Each day, we randomly selected which response direction would yield the high value reward (left or right). After 60 correct trials, contingencies were reversed for 60 trials and then reverted back to the original contingencies.

Figure 4 plots percent choice on free-choice trials (top row), and percent correct (middle row) and reaction time (bottom row) on forced-choice trials, averaged over the first two and last two test days for lesions and controls. A multi-factor ANOVA on number of free choices made to high and low valued outcomes produced significant ($p < 0.05$) interactions of lesion (control and lesion), block (1–3), value manipulation (delay vs size) and day (1–4). Post-hoc t tests revealed that controls ($n = 5$) chose short delay and large reward over long delay and small reward trials, respectively (Figure 4A; control). This was significant in all three trial blocks for the delay manipulation and two out of the three blocks for the size manipulation. For lesions ($n = 5$), there was no difference between the selection of high and low value outcomes in any trial block during the first two days for either delay (Figure 4A; lesion) or size (Figure 4B; lesion) manipulation. Importantly, this was true during the first block of trials for both size and delay suggesting that lesioned rats genuinely had issues with selecting the more valuable option, instead of simply having difficultly reversing contingencies. These results add to the growing body of literature that has proposed different roles for NAc during discounting (Cardinal et al., 2001; Acheson et al., 2006; Floresco et al., 2008; Roesch and Bryden, 2011).

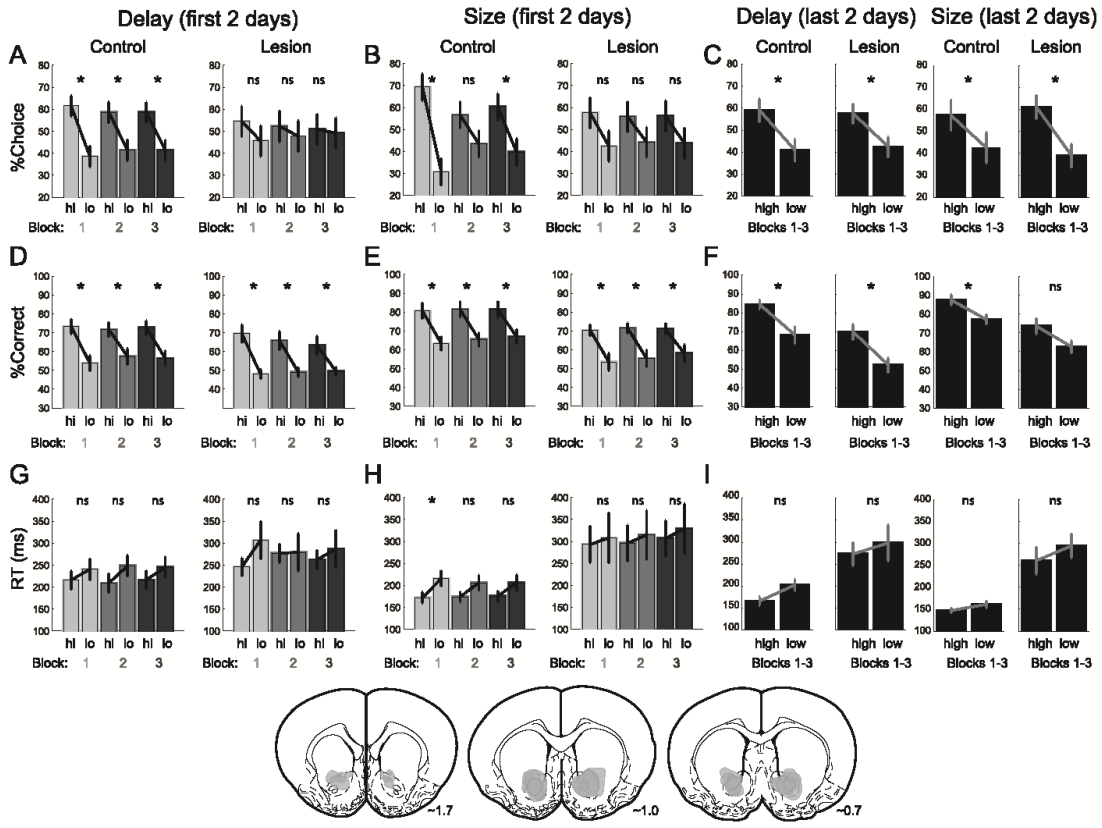

Figure 4. VS lesions caused temporary impairments of reward-guided decision-making. Percent choice on free-choice trials (top row), percent correct scores on forced-choice trials (second row) and reaction time on forced-choice trials (bottom row; port exit minus odor offset) for controls (n = 5) and lesions (n = 5) during the first and last 2 days of testing. Each day rats performed 3 trial blocks of either size or delay across 4 days for each manipulation. During the first 2 days of testing scores were broken down by the three blocks to demonstrate that VS lesions impaired all three trial blocks. Gray areas mark lesions for each animal. Shown are representative slices at 1.7, 1.0 and 0.7 anterior to bregma taken from Paxinos and Watson (1997). High (hi) = short and large. Low (lo) = long and small. Asterisks indicate planned comparisons revealing statistically significant differences (t test, $p<0.05$). Error bars indicate SEM.

Although there was an initial impact of the lesion on free choice behavior, by the last two testing days for delay and size manipulations, both groups were significantly choosing the high over low value outcomes (t test; $p < 0.05$). The average over blocks during the last two days is shown in Figure 4C. Unlike performance on the first two days of testing (Figure 4A and B), lesioned rats significantly chose higher value reward in every block during the last two days (t test; $p < 0.05$).

Overall, lesioned animals were slower and less accurate on forced-choice trials (Figure 4D–I). The multi-factor ANOVA on percent correct and reaction time data produced main effects of lesion and interactions with day and value manipulation (ANOVA; $p < 0.05$). Post-hoc t tests reveal that percent correct scores were significantly better for both value manipulations across blocks in controls and lesions (Figure 4D and E, t test; $p < 0.05$). Finally, in most cases the individual block comparisons between high and low value reaction times were not significantly different, but in all cases rats from both groups tended to be faster under high value conditions (Figure 4G–I; high versus low).

We conclude that NAc lesions do impair the ability to make reward-guided decisions on this task, but this impairment is transitory, possibly reflecting the increased selectivity for odor stimuli and response direction observed in DLS after unilateral lesions.

DISCUSSION

The transition from goal-directed to S-R driven behavior is thought to depend on connectivity between VS and DS via midbrain DA neurons (Barto, 1995; Houk et al., 1995; Sutton and Barto, 1998; Haber et al., 2000; Joel et al., 2002; Redish, 2004b; Niv and Schoenbaum, 2008; Takahashi et al., 2009; Van der Meer and Redish, 2011). Chronic drug use amplifies the strength of this transition, shifting the balance of encoding from VS to DS (Everitt et al., 2001; Everitt and Robbins, 2005; Vanderschuren et al., 2005; Takahashi et al., 2007, 2009; Belin et al., 2009; Corbit et al., 2012).

Here we show that lesions to the NAc region of VS alone, independent of drug use, can enhance stimulus and response selectivity in DLS. It is important to note that we are not suggesting that NAc lesions mimic structural changes that occur in addiction. However, this procedure does allow us to eliminate value signals generated by NAc that we know are diminished after chronic cocaine use (Takahashi et al., 2007). The advantage of this study is that we were able to examine changes in DLS selectivity independent from other changes that might occur during acute or chronic drug use (e.g., receptor availability; disruption of other areas).

One interpretation of our data is that encoding in DLS is heightened to compensate for the loss of NAc function, as has been described for VS after DS lesions (Nishizawa et al., 2012). During performance of our task, normal animals likely base decisions on a mixture of outcome expectancies and S-R contingencies. Without NAc the rats likely depend more heavily on S-R encoding in DLS during decision-making, which allows behavior to recover over several days.

The mechanism by which NAc alters encoding in DLS after lesions and chronic drug use might reflect abnormal DA signals (Haber et al., 2000; Ikemoto, 2007; Takahashi et al., 2009). Indeed, with extended drug self-administration there is elevated DA efflux in DS (Ito et al., 2000, 2002, 2004), which might lead to an excessive stamping in of associations between stimuli and responses. However, recent work shows that increases in DA release in DS during extended periods of self-administration is dependent on NAc (Willuhn et al., 2012). That is, lesions to NAc abolish DA release to cues that predict cocaine. This suggests that lesions of NAc do not increase cue-evoked DA release in DLS, at least in the context of drug self-administration. Whether or not this is true during unexpected reward delivery in a non-drug setting is unknown. Nevertheless, reduced DA signals arriving to DLS might evoke compensatory mechanisms that elevate processing of stimuli and responses in brain areas that signal this information.

Regardless of the underlying mechanism, it is clear that NAc lesions alone increase stimulus and response encoding in DLS. Taken together with our previous findings showing an increase in R-O encoding in DLS after cocaine exposure, this suggests that during the development of addiction, the transition from VS to DS governed behavior due to initial loss of NAc function is not as clear as once thought.

Chapter 5: General Discussion

Summary of Results

The striatum is a highly integrative structure in the basal ganglia involved in the selection and execution of actions with inputs from many cortical and sub-cortical structures (Graybiel et al., 1994; Kreitzer and Malenka, 2008; Matamales et al., 2009). This structure plays an important role in processing emotional, cognitive, and sensorimotor information to generate or suppress actions (Graybiel et al., 1994; Matamales et al., 2009). This is essential in successful navigation of one's environment by either approaching and selecting highly valued outcomes, or suppressing and avoiding negative outcomes. There have been many single-neuron and interference studies showing that the striatum is critical for reward-guided and habitual behaviors (Burton et al., 2015; Bissonette and Roesch, 2016). Research has shown that reward-based learning and decision-making can fall under the control of goal-directed and habit-like processing within striatal sub-regions (Balleine and O'Doherty, 2010).

More specifically, goal-directed behaviors are under the control of value-related response-outcome associations (e.g., "If I do this, I will get a reward"), which fall under the control of ventral regions of striatum such as the NAc and more dorsal medial regions of striatum (DMS) (Balleine and O'Doherty, 2010; Burton et al., 2014a, 2015). On the other side of the spectrum, behaviors related to habitual processing and responding after extended periods of learning are under the control of stimulus-response associations (e.g., "When I see this cue, I automatically respond").

Neural correlates related to these behaviors are found in more lateral regions of dorsal striatum (DLS) (Balleine and O'Doherty, 2010; Burton et al., 2015).

Although the NAc, DMS, and DLS are not directly connected, there is evidence of spiraling circuitry between ventral regions of striatum, the DA system in the midbrain, and dorsal regions of striatum (Haber et al., 2000; Belin and Everitt, 2008; Haber and Knutson, 2009; Everitt and Robbins, 2013; Burton et al., 2015). This communication allows for the propagation of information from more limbic regions (NAc) to associative and sensorimotor regions (DMS and DLS), which allows for behaviors to become more stimulus-driven and efficient over time (Bissonette and Roesch, 2016). This is consistent with the idea that at first behaviors are goal-directed and over time with extended training become more habitual in nature (Balleine and O'Doherty, 2010). Further, this also fits with current theories of addiction as a loss of goal-directed behavioral control governed by more ventral regions of striatum in favor of more habit-like control of behaviors governed by more dorsal regions of striatum (Everitt and Robbins, 2005, 2013, 2016; Hyman et al., 2006; Belin and Everitt, 2008).

Through strong dopaminergic innervation from the midbrain, the striatum and in particular the NAc, is one of the first regions in the brain to be targeted by drugs of abuse, including cocaine (Keiflin and Janak, 2015; Lüscher, 2016). Indeed, synaptic plasticity (LTP) in the NAc has been observed after just one injection of cocaine (Ungless et al., 2001). Long-term changes of chronic drug abuse have also been observed—including synaptic plasticity, cellular epigenetic mechanisms, significant spinogenesis, and elevated levels of nuclear proteins even after periods of abstinence

(Conrad et al., 2008; Wolf and Ferrario, 2010; Lee et al., 2013; Nestler, 2013). However, how encoding in single neurons changes after exposure to cocaine is still not fully understood. The first objective of my research was to determine if previous cocaine self-administration impacted behavior and signaling in NAc during performance of our odor-guided decision-making task where we manipulated the value of reward by changing the delay to or size of reward across a series of blocks of trials (Chapter 2).

In Chapter 2, we showed that rats previously exposed to cocaine exhibited significantly stronger response biases towards higher valued reward towards the end of delay blocks, significantly decreased counts of neurons in the NAc that were responsive to reward and odor epochs, diminished directional and value encoding at the time of reward, and attenuated signaling to longer delayed rewards in NAc single neurons and the population as a whole compared to controls. Cocaine-exposed rats chose shorter delays over longer delays significantly more frequently compared to controls on free-choice trials by the end of the trial blocks (Chapter 2, Figure 1B). Furthermore, rats exposed to cocaine were significantly worse and faster on forced-choice trials compared to controls (Chapter 2, Figure 1C-D). Overall, consistent with previous findings, we concluded that cocaine made rats more impulsive and sensitive to higher valued outcomes in delay blocks in our task.

After cocaine exposure, we observed significantly fewer numbers of neurons in the NAc responsive to both odor and reward epochs in the task compared to controls. In reward-responsive neurons that remained, we observed significantly diminished value and directional encoding in rats exposed to cocaine (Chapter 2,

Figure 2D-E and Figure 3). Control rats showed typical value-encoding in the preferred direction (Chapter 2, Figure 2B-C and Figure 3), with significantly increased activity and significantly more responsive neurons for short over long-delay rewards and for large over small rewards during the reward period. Furthermore, we found that activity in the NAc over longer delays was attenuated in cocaine-exposed rats and remained stable across longer delays in controls (Chapter 2, Figure 2A and Figure 4). We conclude that cocaine made rats more impulsive in the task, and that changes reflected in the NAc were consistent with the impact on behavior (discussed in more detail below).

Through the spiraling circuitry described above and the current theories of drug abuse claiming loss of goal-directed control in favor of habitual control of behavior, one would hypothesize that over time the dorsal striatum is also affected by drugs of abuse. The next objective of my research was to determine if previous cocaine self-administration impacted signaling in DLS during performance of our odor-guided decision-making task (Chapter 3).

In Chapter 3, we showed that rats previously exposed to cocaine exhibited stronger behavioral responses toward higher valued reward in the odor-guided decision-making task and an over-representation of a type of R-O encoding (action-outcome signals) at both the single neuron and population level. Rats that had been previously exposed to cocaine chose higher valued outcomes (shorter delays and larger rewards) on free-choice trials at significantly higher rates towards the end of the block compared to control rats (Chapter 3, Figure 1B). In addition, cocaine-exposed rats performed significantly better and faster on high-value forced-choice

trials towards the end of the delay blocks and significantly better on high-value forced-choice trials towards the end of size blocks in comparison to controls (Chapter 3, Figure 1C-D). Overall, we concluded that previous cocaine exposure made rats more sensitive and impulsive to high value outcomes in the odor-guided decision-making task, which fit with other results showing increased impulsivity in the same (Chapter 2 and (Roesch et al., 2007b)) and other delay discounting tasks (Simon et al., 2007; Dandy and Gatch, 2009; Setlow et al., 2009; Mendez et al., 2010).

In the DLS of cocaine-exposed rats, there was an over-representation of action-outcome signals at both the single neuron and population level (Chapter 3, Figure 2A, 3B, and 4C-D). In controls, we observed a common type of response-outcome encoding that we termed "chosen-outcome" encoding—an increase in activity for the direction the rat was to select and the outcome that was predicted by that selection (Chapter 3, Figure 3A, and 4A-B). However, in cocaine rats, population firing in DLS did not reflect the outcome of the action that was about to occur (chosen-outcome encoding), but instead activity represented the location of the preferred outcome within the context of a particular block of trials (action-outcome encoding; Chapter 3, Figure 3B, and 4C-D). This increase in action-outcome encoding, a type of R-O encoding related to more goal-directed behaviors, was surprising to us and it is important to note we did not observe any increases in S-R encoding (related to habitual processing and responding) after cocaine exposure.

Current theories suggest that the transition from goal-directed to S-R driven behavior depends on the connectivity between NAc and DLS via the spiraling circuitry with midbrain DA neurons. Further, it is thought that chronic drug use

amplifies the strength of this transition, shifting the balance of encoding from more ventral regions of striatum to more dorsal regions. Although we did observe diminished directional and value encoding in NAc after cocaine exposure, fitting with theories of addiction being an initial loss of goal-directed neural correlates governed by this region, we did not observe any increases in S-R correlates in DLS after cocaine-exposure. Surprisingly, we observed increased forms of R-O encoding and goal-directed behaviors that remained intact during our task performance. This led us to ask if pharmacological lesions of NAc, irrespective of any drug use, would impact neural activity in DLS during performance of the odor-guided decision-making task (Chapter 4). A loss of NAc function irrespective of any drug use, allowed us to eliminate value signals from NAc and examine changes in DLS independent from other changes that may occur after drug use.

To address this, we performed unilateral excitotoxic lesions of NAc and implanted electrodes in the same hemisphere in DLS. We observed significantly more and fewer increasing- and decreasing-type neurons in the DLS of lesioned rats, respectively (Chapter 4, Figure 2D). In addition, we observed an overall increase in population activity in response to odor cues in lesioned rats (Chapter 4, Figure 2C). When we looked at how individual neurons were modulated by expected outcome, response direction, and stimulus type there were increased numbers of increasing-type neurons that were responsive to the contralateral response direction and increased numbers of decreasing-type neurons for stimulus-type in lesioned rats (Chapter 4, Figure 3E-F). We concluded that after NAc lesions, there was an increase in stimulus and response encoding in DLS. Furthermore, we found that NAc was

necessary for normal performance of our odor-guided decision-making task: bilateral lesions of NAc caused a transitory decrease in value-based decision-making during initial days of testing (Chapter 4, Figure 4). This was temporary, as the last days of testing showed comparable behavior between lesions and controls in the task.

To summarize the results in this dissertation, we found that after previous cocaine self-administration there were behavioral and neural changes observed in NAc and DLS. After cocaine exposure, we saw an increase in impulsive behavior and an increased bias towards higher value rewards. Further, in cocaine rats the NAc showed fewer odor- and reward-responsive neurons and in those reward-responsive neurons that remained we observed diminished value and directional encoding. Cocaine rats also showed attenuated activity in NAc for longer delays compared to controls. In DLS, surprisingly, we did not observe any increases in S-R encoding, but an increase in action-outcome encoding after cocaine exposure. DLS population activity in cocaine rats better reflected the preferred outcome in a particular block of trials, regardless of whether or not the animal moved in the direction to obtain that outcome. Finally, irrespective of any drug-use we showed that NAc lesions increased S-R encoding in DLS during performance of the odor-guided decision-making task that was dependent on the normal function of NAc.

Overall, we conclude that drug exposure is not as simple as silencing the NAc and enhancing S-R processing under control of DLS as we observed after lesions. These results challenge common theories of drug addiction that state that enhanced habitual responding is at the cost of goal-directed control of behavior, leading to maladaptive decision-making commonly seen in this disorder. Indeed, we found

evidence that cocaine-exposure did not alter goal-directed decision-making—cocaine rats were actually better at choosing higher valued outcomes compared to controls. We did observe a loss in goal-directed behavior after bilateral lesions to NAc, but this was transitory and behavior returned to levels seen in controls. Next, I will discuss the impact of cocaine and lesions on behavior, and potential explanations of how these results fit together.

The impact of cocaine and lesions on behavior in the odor-guided decision-making task

Delay discounting tasks are often used to measure impulsivity, which is known to be a problem in addicts performing delay discounting, gambling, and reversal learning tasks (Heil et al., 2006; Reynolds, 2006; Robbins et al., 2008; Mackillop, 2013; Volkow and Baler, 2015; Everitt and Robbins, 2016). The rate at which one biases behavior towards smaller, more immediate rewards over larger, more delayed rewards is used as a measure of impulsivity. We have shown that previous cocaine exposure increases impulsivity towards high-value rewards in the odor-guided decision-making task, particularly during delay blocks. In both studies (Chapter 2 and 3), cocaine-exposed rats chose shorter delays over longer delays significantly more than controls by the end of the trial block. This is consistent with other findings showing that rats exposed to cocaine are quicker to abandon longer delays in favor of more immediate rewards (Roesch et al., 2007b; Simon et al., 2007; Dandy and Gatch, 2009; Setlow et al., 2009; Mendez et al., 2010) and it fits with

other studies showing addicts are more impulsive (Heil et al., 2006; Reynolds, 2006; Mackillop, 2013; Volkow and Baler, 2015).

In Chapter 3, we also observed increased impulsivity on size blocks in that cocaine rats chose larger reward significantly more often than small reward by the end of trial blocks compared to cocaine. Furthermore, in this group of rats we observed a higher sensitivity to high-value outcomes on forced-choice trials in that these rats were significantly better and faster on high-value outcomes by the end of both delay and size blocks compared to controls. This increased impulsivity and sensitivity to high-value outcomes (short delays and large rewards) was consistent with previous findings using the same task (Roesch et al., 2007b). It is worth noting that we did not observe this in the other set of animals (Chapter 2), but regardless of why this may be, we clearly see increased sensitivity to delay manipulations across both studies. Further, the majority of work examining the impact of drugs of abuse on delay-discounting have shown effects on delay but not size (Simon et al., 2007; Mendez et al., 2010).

Overall, we conclude that previous cocaine self-administration increases impulsivity in our odor-guided decision-making task. When considering underlying behavioral control and how cocaine impacts goal-directed and habitual responding, we would expect cocaine rats to be inflexible in their choices and show a loss of goal-directed decision-making. However, we observe cocaine rats actually choosing higher value outcomes more often when freely able to choose and better performance on forced-choice trials, albeit faster, compared to controls. This suggests that they are

still able to make goal-directed decisions in our task, contrary to current addiction theories suggesting a loss of goal-directed behavioral control.

In Chapter 3, we did observe some evidence of inflexibility in cocaine rats—they were actually slower to reverse contingencies early in size blocks, choosing small reward more often than large reward in the first ten free-choice trials (Chapter 3, Figure 1B). This fits with the theory of habitual control of behavior and being inflexible in decision-making, however as the block went on cocaine rats biased their behavior towards larger rewards. Overall, this inflexibility was not seen on the majority of blocks in either study and we conclude that cocaine rats are still able to make goal-directed decisions as contingencies change in the task.

Previous research has shown that goal-directed behaviors are governed by ventral regions of striatum, where R-O correlates during decision-making emerge. To determine if NAc, a region of VS, was necessary for normal decision-making in our task we performed bilateral lesions to this brain region and then looked at task performance. Indeed, we did find an initial loss of goal-directed behavioral control in the first days of testing after lesions. Lesioned rats chose equally between high and low value rewards during the first two days of testing. This effect was transitory, as the last two days of testing showed that lesioned rats significantly chose the higher value reward more often than low value reward across all blocks of trials. The initial loss of goal-directed decision-making fits with current theories of addiction. It is important to consider that this effect was temporary in nature, we suspect that other brain regions (such as DLS) were taking over and allowing for increased performance in the task by the end of testing.

When taken together, previous cocaine self-administration left goal-directed mechanisms of behavioral control intact while lesions caused an initial loss of goal-directed decisions in our task. Our lesion study fits with current theories of addiction as NAc being one of the first areas to become damaged or affected by drugs of abuse, which may contribute to a loss of goal-directed behaviors and R-O correlates in favor of more habitual and S-R driven behaviors and neural correlates governing decision-making. However, our cocaine self-administration studies do not fit with this theory and even challenge it in many ways. Cocaine rats were still able to make goal-directed decisions, and in fact were better than controls on some behavioral measures.

We speculate that cocaine alters signaling in intricate ways beyond just eliminating R-O correlates and that the subsequent effects on behavior manifest in more complex ways than simply losing goal-directed decision-making mechanisms. Our task often changes contingencies, with value associations changing in each block. This could prevent S-R driven behaviors from developing and even encourage goal-directed behavioral control, but we feel that the complexity of our task is a better representation of the everyday complex decisions one faces. This fits with others who have shown that under similar circumstances exposure to drugs of abuse does indeed leave goal-directed mechanisms intact and in some cases, enhanced (Phillips and Vugler, 2011; Son et al., 2011; Halbout et al., 2016).

More specifically, Phillips and Vugler (2011) and Son et al. (2011) used contingency degradation paradigms, where contingencies between lever-presses and reinforcers were manipulated, to ask whether action-outcome relationships would change after stimulant exposure. They showed that after stimulant exposure

(amphetamine/methamphetamine), sensitized animals reduced responding during the contingency degradation phase where the reinforcer was no longer paired with lever-presses at a similar rate compared to controls (Phillips and Vugler, 2011; Son et al., 2011). Further, they showed sensitized animals maintained responding on non-degraded contingencies while controls reduced this response over time (Phillips and Vugler, 2011).

In another recent study, Halbout and colleagues (2016) showed similar results on a contingency degradation paradigm after cocaine exposure. When rats were trained on two different action-outcome contingencies, rats exposed to cocaine decreased responding on degraded outcomes and actually altered their behavior more quickly compared to controls (Halbout et al., 2016). Even though we did not devalue outcomes in our task, our findings are consistent with the results from these studies in that we do not see enhanced S-R encoding (Chapter 3), do see increased action-outcome correlates (Chapter 3), and do see faster behavioral adjustments in cocaine-exposed rats when contingencies change (Chapter 2 and 3).

Past research supporting the idea that chronic drug use impairs goal-directed mechanisms governed by R-O encoding while enhancing habitual behaviors governed by S-R encoding have focused on using paradigms where there is a singular response or no real choice at all (Pavlovian) (Robbins and Everitt, 1999; Everitt et al., 2001; Everitt and Robbins, 2005; Schoenbaum and Setlow, 2005; Vanderschuren et al., 2005; Everitt and Robbins, 2016; Nelson and Killcross, 2006, 2013; Nordquist et al., 2007; Ostlund and Balleine, 2008; Redish et al., 2008; Hogarth et al., 2013; LeBlanc et al., 2013; Corbit et al., 2014; Lucantonio et al., 2014; Schmitzer-Torbert et al.,

2015). However, these other studies (including the studies here, Chapter 2-3) using more complex tasks that offer choices between different rewards and situations may better reflect everyday decision-making and evoke multiple neural representations beyond R-O and S-R correlates impacted by drugs of abuse.

Connecting the behavioral impact of cocaine and lesions to changes in neural signaling in the striatum

In Chapter 2, we described several alterations in NAc that might contribute to abnormal task performance that we observed in this group of rats. First, we observed an overall reduction in the number of cells responsive to both odor cues and reward delivery. Second, in those cells that remained reward-responsive, we observed a reduction in directional and value-encoding. Fewer neurons and weaker encoding during important time periods in the task might underlie drug-induced changes in behavior including worse performance on forced-choice trials. If NAc is no longer providing accurate feedback about the location and value of upcoming outcomes, this could lead to altered cue associations. This could manifest as a decrease in performance on trials that depend upon accurate cue associations in responding in the correct direction to obtain reward (forced-choice trials), which we do observe.

Additionally, neural firing in the NAc to delayed rewards was weaker in rats that had self-administered cocaine. This is of particular note, as cocaine rats were more sensitive to longer delayed rewards as evidenced by faster reaction times and better choice behavior in favor of the more immediate reward compared to controls. If the NAc cannot accurately maintain neural activity across delays, this could disrupt

behavior on forced-choice trials where they are forced to remain in the well until reward delivery. If they exit the fluid-well before reward delivery, this would result in an incorrect response and drive down overall percent correct on these trial types.

Next, we showed that population firing in DLS in cocaine-exposed rats failed to modify predictions based on the action selected, but instead showed an increase in neural activity reflecting the location of the preferred outcome regardless of subsequent movements (Chapter 3). This activity after cocaine exposure was less-stimulus bound, in that within a particular block of trials neurons fired whether the cue was a free- or forced-choice odor *and* activity emerged before odor onset. An enhancement of these signals would cause fast, biased decisions after cocaine exposure. These altered neural correlates likely bias behavior successfully on free-choice trials, while diminishing performance on forced-choice trials.

When considering the indirect connections between the NAc and the DLS and the propagation of information being sent along this pathway, we speculate that the behavioral and neural changes seen after cocaine exposure in these studies is tightly linked to changes in DA signaling. The spiraling DA circuit between midbrain regions and striatal targets is known to be affected by drug-seeking and drug-taking (Haber et al., 2000; Belin and Everitt, 2008; Haber and Knutson, 2009; Everitt and Robbins, 2013). Reward prediction errors (reward experienced minus reward expected) are critical in learning about the current environment and are signaled by DA neurons (Schultz, 2006). It has recently been shown that reward prediction errors signaled by dopaminergic neurons in the VTA depend upon accurate signaling from

the NAc, specifically in relation to the timing of rewarded outcomes (Takahashi et al., 2016) and that this is essential to developing neural selectivity to cues in DA neurons and neural signaling in the NAc (Nicola et al., 2005; Takahashi and Schoenbaum, 2016).

Recently, Saddoris and colleagues showed evidence of altered DA signaling in the NAc after withdrawal from self-administering cocaine for two weeks. They found that DA release in the NAc does not distinguish between reward-predicting cues and was abnormally biased towards reward delivery after cocaine exposure (Saddoris et al., 2016b). A bias toward reward delivery after cocaine exposure could reflect an abnormal increase in positive reward prediction errors, which could bias behavior towards more valuable reward during choice performance on our task.

Taken together, these findings fit with our overall result of fewer cue-responsive neurons in the NAc after cocaine exposure, which ultimately could impact processing of upcoming cues, rewards, and associated actions in downstream areas such as DLS. Consistent with this idea, we showed that lesions to NAc (Chapter 4) and exposure to cocaine impacts encoding in DLS (Chapter 3) albeit, in different ways. After NAc lesions, neural correlates related to stimulus and response encoding in DLS were enhanced (Chapter 4). After cocaine self-administration, stimulus-response correlates were unaltered, whereas correlates related to response-outcome encoding were enhanced and divorced from actions (Chapter 3). This lack of processing of direction and value in DLS might have resulted from impaired encoding in NAc.

The impact of cocaine in the reward-circuit and how our results fit within the circuit

Even though the striatum, and in particular the NAc, is hypothesized to be one of the first regions to be affected by drugs of abuse, it is important to consider other brain regions within the reward circuit that could also be affected early on by addiction and subsequently alter neural signaling within the reward circuit. As I mentioned in Chapter 1, the OFC and ABL are two areas in the reward circuit that are critical in reward-guided decision-making. The OFC signals expected outcomes and depends on ABL to encode cue-outcome associations to develop these predictions (Schoenbaum et al., 1998, 2003b; Tremblay and Schultz, 1999; O'Doherty et al., 2002; Gottfried et al., 2003; Hikosaka and Watanabe, 2004; Roesch and Olson, 2004, 2005; Blair et al., 2006; Roesch et al., 2006; Feierstein et al., 2006; Roberts, 2006; Dolan, 2007; Furuyashiki et al., 2008). These regions also show structural and functional changes after exposure to drugs of abuse in monkeys and rats, and show damage in addicts in human imaging studies (Jones and Mishkin, 1972; Rolls et al., 1994; Dias et al., 1996; Meunier et al., 1997; Bohn et al., 2003; McAlonan and Brown, 2003; Schoenbaum et al., 2003a; Izquierdo et al., 2004; Kim and Ragozzino, 2005; Chudasama et al., 2007).

To the best of our knowledge, the studies in this dissertation provide novel evidence of altered neural signaling in striatal regions during performance of a complex behavioral decision-making task that changes contingencies often. Previously in the literature, a common way of examining executive cognitive functions after drug exposure has been to look at measures of cognitive flexibility, which refers to the ability to adapt one's cognitive representations and subsequent

behaviors to changing conditions in the environment (Stalnaker et al., 2009). We know that addicts have difficulties in changing responses to drug-seeking and -taking, particularly when faced with stimuli that are associated with the drug itself. Evidence from human studies shows that human cocaine addicts are impaired in gambling and reversal learning tasks that both probe cognitive flexibility (Rogers et al., 1999; Grant et al., 2000; Rogers and Robbins, 2001; Bechara and Dolan, 2002; Jentsch et al., 2002; Fillmore and Rush, 2006; Ersche et al., 2008).

In particular, reversal learning tasks have been used in rodents to examine cognitive flexibility after exposure to cocaine (Schoenbaum et al., 2004; Calu et al., 2007; Stalnaker et al., 2009). Similar to our task, this reversal task used odor cues to either predict a liquid sucrose reward or quinine punishment. After a series of trials, these odor contingencies switched and the rate at which the animals learned to respond to initial contingencies and then the reversal was looked at in conjunction with neural correlates in OFC, ABL, and striatum. Rats exposed to cocaine, either passively or through self-administration, are abnormally slow to learn reversals even though they learn initial contingencies at similar rates compared to controls (Schoenbaum et al., 2004; Calu et al., 2007; Stalnaker et al., 2009).

The importance of the OFC and ABL in reversal learning and cognitive flexibility is its ability to signal the value of an expected outcome and properly encode cue-outcome associations, respectively (Fuster and Uyeda, 1971; Cain and Bindra, 1972; Sanghera et al., 1979; Quirk et al., 1995; Nishijo et al., 1997; Schoenbaum et al., 1998, 1999, 2003b; Tremblay and Schultz, 1999; Maren, 2000; O'Doherty et al., 2002; Toyomitsu et al., 2002; Gottfried et al., 2003; Roesch and

Olson, 2004, 2005; Hikosaka and Watanabe, 2004; Saddoris et al., 2005; Blair et al., 2006; Roberts, 2006; Roesch et al., 2006; Feierstein et al., 2006; Dolan, 2007; Furuyashiki et al., 2008). This would contribute to flexible behavior by allowing the recognition after reversal that the value of the outcome no longer matches that of the predicting cue. Indeed, animals with OFC lesions are impaired in learning reversals (Saddoris et al., 2005; Stalnaker et al., 2007a). Furthermore, it has been shown that damage to OFC alters ability of ABL neurons to signal cue-outcome associations necessary for normal performance on reversals (Saddoris et al., 2005). This suggests that OFC is necessary for facilitation of flexible encoding in ABL through its ability to signal outcome expectancies. Interestingly, lesions to both OFC and ABL had no effect on reversal learning, suggesting that other regions (such as striatum) support reversal learning (Jog et al., 1999; Setlow et al., 2003; Schmitzer-Torbert and Redish, 2004; Yin et al., 2004; Yin and Knowlton, 2006; Takahashi et al., 2007). Lesions to striatum confirm this role, as performance is impaired after lesions to VS and DS (Ferry et al., 2000; Ragozzino et al., 2002).

After cocaine exposure, neurons in OFC failed to develop cue selectivity that predicted the outcome where control rats showed robust cue selectivity developing for both sucrose and quinine and to the cues predicting these outcomes (Stalnaker et al., 2006, 2009). Further, ABL neurons in cocaine rats showed inflexible coding during reversal tasks, failing to reverse cue-selectivity after reversals (Stalnaker et al., 2007b, 2009). ABL neurons in control rats developed selectivity to each odor cue and this reversed after reversals, suggesting a flexible encoding mechanism in this region (Schoenbaum et al., 1999; Saddoris et al., 2005). The striatum also showed subtle

changes in reversal performance after cocaine exposure (Takahashi et al., 2007). Cue-selectivity was less flexible in the VS of cocaine rats, and there was an overall decrease in cue encoding. The DLS of cocaine rats did not show any significant differences compared to controls, suggesting that cognitive inflexibility after cocaine may arise more from functional changes in the OFC and ABL compared to striatum.

These results fit with my results showing that overall there were significantly lower numbers of neurons responsive to the odor cue in the NAc or VS in our task. Although we do not have recordings from OFC or ABL after cocaine exposure in the task in this dissertation, we can speculate from the reversal learning results about the impact on the circuit overall. If OFC loses the ability to signal expected outcome value and ABL loses the ability to flexibly encode cue associations, we would expect this to impact downstream areas such as striatum. Specifically, I would expect NAc to lose the ability to signal predicted outcomes based on cues, which has been shown in NAc after OFC lesions (Cooch et al., 2015). Indeed, we observe significantly fewer odor- and reward-responsive cues and diminished value encoding in reward-responsive neurons that remain in the NAc of cocaine rats.

The loss of important and relevant cue and outcome information would impact signaling in even further downstream areas, such as the DLS, that is more closely related to control of behavior. Indeed, I show evidence that DLS is impacted both by lesions to NAc and cocaine exposure. Loss of NAc increased S-R encoding in DLS. However, cocaine increased R-O associations in the DLS that were removed from selected actions, suggesting that the DLS was signaling the location of a preferred outcome regardless of whether or not actions were taken to obtain it. Taken together,

from these studies focusing on decision-making tasks, whether reversal learning or our odor-guided delay/size task, it is evident that cocaine alters many regions throughout the reward circuit that are critical for normal reward-guided processing and decision-making.

Future directions

Further work needs to be done to specifically determine the role DA plays in decision-making after drug exposure. Reward prediction errors signaled by DA neurons are thought to modify decision-making by updating neural representations of reward value in NAc. Past research has shown that sub-second DA release increases and decreases in response to events that are better and worse than predicted, respectively. It is still unknown how prediction error signaling by the DA system is disrupted after chronic exposure to drugs of abuse and how this may lead to abnormal decision-making. To address this issue, future studies should focus on DA release patterns in complex decision-making tasks such as the one used in these studies after cocaine self-administration.

We have shown that rats bias their behavior towards reward that is more immediate or larger than expected (positive reward prediction error, increase in DA firing) and away from rewards that are delayed or smaller than expected (negative reward prediction error, decrease in DA firing). We, and others, have shown that rats chronically exposed to cocaine exhibit abnormal behavior during performance of tasks that manipulate reward size and delays to reward (i.e. delay discounting; intertemporal choice). An underlying hypothesis is that that these changes in

behavior are due to cocaine-induced disruption of reward prediction error related DA release in VS. All aspects of prediction error encoding could be examined with fast-scan cyclic voltammetry, a recording technique that allows high temporal precision in recordings of neurotransmitter release. It has been recently shown that DA release specifically during cue sampling is severely disrupted after cocaine exposure (Saddoris et al., 2016b), and we show that significantly fewer neurons in the NAc were responsive to odor cues (Chapter 2). Further studies could be done to rescue disrupted prediction error signals via optogenetic manipulation in TH-CRE rats by stimulating DA release in NAc at precise time points during performance of our task.

More specifically, after training on the task rats would be implanted with catheters for cocaine self-administration, chronic electrodes in NAc and DLS, and bilateral virus/optical fibers in the VTA to stimulate DA neurons. There would be four main groups: two groups will self-administer cocaine and two will self-administer sucrose pellets. Within the cocaine and control group, you would have a viral expressing and non-expressing group to serve as a control for optogenetic stimulation. After recovery from surgery, rats will go through the same self-administration protocol and withdrawal time described in this dissertation.

Once recordings start, a blue laser would be used to activate the virus in DA cells specifically during odor cue sampling on forced-choice trials. This activation of DA neurons would mimic naturally-occurring DA transients found to be altered at the time of cues in the NAc. Bilateral implantation of optical fibers would allow you to stimulate either bilaterally or unilaterally, to look at overall behavior or neural signaling in the same hemisphere that was stimulated, respectively. Specifically, after

activation, you could examine neural activity from recording electrodes in the NAc and DLS and behavior to determine if neural signals and behavior improved.

I would predict that repairing these DA signals in NAc after cocaine-exposure would improve performance on forced-choice trials in our task and make rats less sensitive to manipulations of size and delay on free-choice trials. I would also expect that rescued signals in NAc would affect downstream signaling in DLS and eliminate enhanced R-O correlates seen in our study (Chapter 3). The ultimate goal in these future studies is to improve our understanding of how neural signals in learning and decision-making circuits are altered after drug exposure. A more complete understanding of the normal and abnormal function of these circuits is critical in aiding the development of more effective treatments for those suffering from addiction.

REFERENCES

Acheson A, Farrar AM, Patak M, Hausknecht KA, Kieres AK, Choi S, de Wit H, Richards JB (2006) Nucleus accumbens lesions decrease sensitivity to rapid changes in the delay to reinforcement. Behav Brain Res 173:217–228.

Ainslie G (1974) Impulse control in pigeons. J Exp Anal Behav 21:485–489 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1333221&tool=pmcentrez&rendertype=abstract.

Albin RL, Young AB, Penney JB (1989) The functional anatomy of basal ganglia disorders. Trends Neurosci 12:366–375.

Apicella P, Ljungberg T, Scarnati E, Schultz W (1991) Responses to reward in monkey dorsal and ventral striatum. Exp Brain Res 85:491–500.

Ballard K, Knutson B (2009) Dissociable neural representations of future reward magnitude and delay during temporal discounting. Neuroimage 45:143–150.

Balleine BW, Delgado MR, Hikosaka O (2007) The Role of the Dorsal Striatum in Reward and Decision-Making. J Neurosci 27:8161–8165 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1554-07.2007.

Balleine BW, O'Doherty JP (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35:48–69.

Barto AG (1995) Adaptive Critics and the Basal Ganglia. In: Models of Information Processing in the Basal Ganglia, pp 215–232 Available at: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6283098.

Bechara A, Dolan S (2002) Decision-making and addiction (part II): Myopia for the future or hypersensitivity to reward? Neuropsychologia 40:1690–1705.

Belin D, Everitt BJ (2008) Cocaine Seeking Habits Depend upon Dopamine-Dependent Serial Connectivity Linking the Ventral with the Dorsal Striatum. Neuron 57:432–441.

Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ (2009) Parallel and interactive learning processes within the basal ganglia: Relevance for the understanding of addiction. Behav Brain Res 199:89–102.

Bezzina G, Cheung THC, Asgari K, Hampson CL, Body S, Bradshaw CM, Szabadi E, Deakin JFW, Anderson IM (2007) Effects of quinolinic acid-induced lesions of the nucleus accumbens core on inter-temporal choice: A quantitative analysis. Psychopharmacology (Berl) 195:71–84.

Bissonette GB, Burton AC, Gentry RN, Goldstein BL, Hearn TN, Barnett BR, Kashtelyan V, Roesch MR (2013) Separate Populations of Neurons in Ventral Striatum Encode Value and Motivation. PLoS One 8.

Bissonette GB, Roesch MR (2016) Neurophysiology of reward-guided behavior: Correlates related to predictions, value, motivation, errors, attention, and action. Curr Top Behav Neurosci 27:199–230.

Björklund A, Dunnett SB (2007) Dopamine neuron systems in the brain: an update. Trends Neurosci 30:194–202.

Blair K, Marsh AA, Morton J, Vythilingam M, Jones M, Mondillo K, Pine DC, Drevets WC, Blair JR (2006) Choosing the Lesser of Two Evils, the Better of Two Goods: Specifying the Roles of Ventromedial Prefrontal Cortex and Dorsal

Anterior Cingulate in Object Choice. J Neurosci 26:11379–11386 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1640-06.2006.

Bohn I, Giertler C, Hauber W (2003) Orbital prefrontal cortex and guidance of instrumental behaviour in rats under reversal conditions. Behav Brain Res 143:49–56.

Bowman EM, Aigner TG, Richmond BJ (1996) Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. J Neurophysiol 75:1061–1073 Available at: http://www.ncbi.nlm.nih.gov/pubmed/8867118.

Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010) Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. Neuron 68:815–834.

Brown VJ, Bowman EM (1995) Discriminative Cues Indicating Reward Magnitude Continue to Determine Reaction Time of Rats Following Lesions of the Nucleus Accumbens. Eur J Neurosci 7:2479–2485.

Bryden DW, Burton AC, Kashtelyan V, Barnett BR, Roesch MR (2012) Response inhibition signals and miscoding of direction in dorsomedial striatum. Front Integr Neurosci 6:1–15 Available at: http://www.frontiersin.org/Integrative_Neuroscience/10.3389/fnint.2012.00069/abstract [Accessed September 11, 2012].

Bryden DW, Johnson EE, Diao X, Roesch MR (2011a) Impact of expected value on neural activity in rat substantia nigra pars reticulata. Eur J Neurosci 33:2308–2317.

Bryden DW, Johnson EE, Diao X, Roesch MR (2011b) Impact of expected value on neural activity in rat substantia nigra pars reticulata. Eur J Neurosci 33:2308–

2317.

Bryden DW, Johnson EE, Tobia SC, Kashtelyan V, Roesch MR (2011c) Attention for learning signals in anterior cingulate cortex. J Neurosci 31:18266–18274 Available at: http://www.jneurosci.org/content/31/50/18266.short.

Burton AC, Bissonette GB, Lichtenberg NT, Kashtelyan V, Roesch MR (2014a) Ventral striatum lesions enhance stimulus and response encoding in dorsal striatum. Biol Psychiatry 75:132–139.

Burton AC, Kashtelyan V, Bryden DW, Roesch MR (2014b) Increased firing to cues that predict low-value reward in the medial orbitofrontal cortex. Cereb Cortex 24:3310–3321.

Burton AC, Nakamura K, Roesch MR (2015) From ventral-medial to dorsal-lateral striatum: Neural correlates of reward-guided decision-making. Neurobiol Learn Mem 117:51–59.

Busetto G, Higley MJ, Sabatini BL (2008) Developmental presence and disappearance of postsynaptically silent synapses on dendritic spines of rat layer 2/3 pyramidal neurons. J Physiol 586:1519–1527 Available at: http://doi.wiley.com/10.1113/jphysiol.2007.149336%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/18202095.

Cai X, Kim S, Lee D (2011) Heterogeneous Coding of Temporally Discounted Values in the Dorsal and Ventral Striatum during Intertemporal Choice. Neuron 69:170–182.

Cain DP, Bindra D (1972) Responses of amygdala single units to odors in the rat. Exp Neurol 35:98–110.

Calu DJ, Roesch MR, Haney RZ, Holland PC, Schoenbaum G (2010) Neural Correlates of Variations in Event Processing during Learning in Central Nucleus of Amygdala. Neuron 68:991–1001.

Calu DJ, Stalnaker TA, Franz TM, Singh T, Shaham Y, Schoenbaum G (2007) Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. Learn Mem 14:325–328 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2396594&tool=pmcentrez&rendertype=abstract.

Cameron CM, Wightman RM, Carelli RM (2016) One month of cocaine abstinence potentiates rapid dopamine signaling in the nucleus accumbens core. Neuropharmacology 111:223–230.

Cardinal RN, Parkinson JA, Hall J, Everitt BJ (2002) Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. Neurosci Biobehav Rev 26:321–352.

Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW, Everitt BJ (2001) Impulsive choice induced in rats by lesions of the nucleus accumbens core. Sci (New York, NY) 292:2499–2501 Available at: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=11375482&retmode=ref&cmd=prlinks.

Cardinal RN, Winstanley CA, Robbins TW, Everitt BJ (2004) Limbic corticostriatal systems and delayed reinforcement. In: Annals of the New York Academy of Sciences, pp 33–50.

Chudasama Y, Kralik JD, Murray EA (2007) Rhesus monkeys with orbital prefrontal cortex lesions can learn to inhibit prepotent responses in the reversed reward contingency task. Cereb Cortex 17:1154–1159.

Colby CR, Whisler K, Steffen C, Nestler EJ, Self DW (2003) Striatal cell type-specific overexpression of DeltaFosB enhances incentive for cocaine. J Neurosci 23:2488–2493 Available at: http://www.jneurosci.org/content/23/6/2488.abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/12657709.

Colwill RM, Rescorla RA (1985) Instrumental responding remains sensitive to reinforcer devaluation after extensive training. J Exp Psychol Anim Behav Process 11:520–536.

Colwill RM, Triola SM (2002) Instrumental responding remains under the control of the consequent outcome after extended training. Behav Processes 57:51–64.

Conrad KL, Tseng KY, Uejima JL, Reimers JM, Heng L-J, Shaham Y, Marinelli M, Wolf ME (2008) Formation of accumbens GluR2-lacking AMPA receptors mediates incubation of cocaine craving. Nature 454:118–121 Available at: http://www.nature.com/doifinder/10.1038/nature06995.

Cooch NK, Stalnaker T a., Wied HM, Bali-Chaudhary S, McDannald M a., Liu T-L, Schoenbaum G (2015) Orbitofrontal lesions eliminate signalling of biological significance in cue-responsive ventral striatal neurons. Nat Commun 6:7195 Available at: http://www.nature.com/doifinder/10.1038/ncomms8195%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/26006060.

Corbit LH, Chieng BC, Balleine BW (2014) Effects of repeated cocaine exposure on habit learning and reversal by N-acetylcysteine. Neuropsychopharmacology 39:1893–1901 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4059898&tool=pmcentrez&rendertype=abstract.

Corbit LH, Nie H, Janak PH (2012) Habitual alcohol seeking: Time course and the contribution of subregions of the dorsal striatum. Biol Psychiatry 72:389–395.

Cousins MS, Atherton A, Turner L, Salamone JD (1996) Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. Behav Brain Res 74:189–197.

Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. J Neurophysiol 89:2823–2838.

Dandy KL, Gatch MB (2009) The effects of chronic cocaine exposure on impulsivity in rats. Behav Pharmacol 20:400–405 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3867254&tool=pmcentrez&rendertype=abstract.

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. Neuron 69:1204–1215.

Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–1711 Available at:

http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&DbFrom=pubmed&Cmd=Link&LinkName=pubmed_pubmed&LinkReadableName=Related Articles&IdsFromResult=16286932&ordinalpos=3&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum.

Delgado MR (2007) Reward-related responses in the human striatum. In: Annals of the New York Academy of Sciences, pp 70–88.

Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380:69–72 Available at: http://dx.doi.org/10.1038/380069a0.

Dickinson A, Smith J, Mirenowicz J (2000) Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. Behav Neurosci 114:468–483 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/0735-7044.114.3.468.

Dolan RJ (2007) The human amygdala and orbital prefrontal cortex in behavioural regulation. Philos Trans R Soc Lond B Biol Sci 362:787–799 Available at: http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&DbFrom=pubmed&Cmd=Link&LinkName=pubmed_pubmed&LinkReadableName=Related Articles&IdsFromResult=17403643&ordinalpos=3&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum.

Ersche KD, Roiser JP, Robbins TW, Sahakian BJ (2008) Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacology (Berl) 197:421–431.

Evenden JL, Ryan CN (1999) The pharmacology of impulsive behaviour in rats VI:

The effects of ethanol and selective serotonergic drugs on response choice with varying delays of reinforcement. Psychopharmacology (Berl) 146:413–421.

Everitt BJ, Dickinson A, Robbins TW (2001) The neuropsychological basis of addictive behaviour. Brain Res Rev 36:129–138.

Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat Neurosci 8:1481–1489 Available at: http://www.ncbi.nlm.nih.gov/pubmed/16251991.

Everitt BJ, Robbins TW (2013) From the ventral to the dorsal striatum: Devolving views of their roles in drug addiction. Neurosci Biobehav Rev 37:1946–1954.

Everitt BJ, Robbins TW (2016) Drug Addiction: Updating Actions to Habits to Compulsions Ten Years On. Annu Rev Psychol 67:150807174122003 Available at: http://www.annualreviews.org/doi/abs/10.1146/annurev-psych-122414-033457.

Feierstein CE, Quirk MC, Uchida N, Sosulski DL, Mainen ZF (2006) Representation of Spatial Goals in Rat Orbitofrontal Cortex. Neuron 51:495–507.

Ferry a T, Lu XC, Price JL (2000) Effects of excitotoxic lesions in the ventral striatopallidal--thalamocortical pathway on odor reversal learning: inability to extinguish an incorrect response. Exp Brain Res 131:320–335.

Fillmore MT, Rush CR (2006) Polydrug abusers display impaired discrimination-reversal learning in a model of behavioural control. J Psychopharmacol 20:24–32.

Floresco SB, St Onge JR, Ghods-Sharifi S, Winstanley C a (2008) Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. Cogn

Affect Behav Neurosci 8:375–389.

Fowler JS, Volkow ND, Kassed CA, Chang L (2007) Imaging the addicted human brain. Sci Pract Perspect 3:4–16 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2851068&tool=pmcentrez&rendertype=abstract.

Frank MJ (2005) Dynamic Dopamine Modulation in the Basal Ganglia: A Neurocomputational Account of Cognitive Deficits in Medicated and Nonmedicated Parkinsonism. J Cogn Neurosci 17:51–72 Available at: http://www.mitpressjournals.org/doi/10.1162/0898929052880093.

Furuyashiki T, Holland PC, Gallagher M (2008) Rat orbitofrontal cortex separately encodes response and outcome information during performance of goal-directed behavior. J Neurosci 28:5127–5138 Available at: http://www.ncbi.nlm.nih.gov/pubmed/18463266%5Cnhttp://www.jneurosci.org/content/28/19/5127.full.pdf.

Fuster JM, Uyeda AA (1971) Reactivity of limbic neurons of the monkey to appetitive and aversive signals. Electroencephalogr Clin Neurophysiol 30:281–293.

Gerfen C, Engber T, Mahan L, Susel Z, Chase T, Monsma F, Sibley D (1990) D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. Science (80-) 250:1429–1432 Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.2147780.

Giertler C, Bohn I, Hauber W (2003) The rat nucleus accumbens is involved in guiding of instrumental responses by stimuli predicting reward magnitude. Eur J

Neurosci 18:1993–1996.

Giertler C, Bohn I, Hauber W (2004) Transient inactivation of the rat nucleus accumbens does not impair guidance of instrumental behaviour by stimuli predicting reward magnitude. Behav Pharmacol 15:55–63 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15075627.

Goldstein BL, Barnett BR, Vasquez G, Tobia SC, Kashtelyan V, Burton AC, Bryden DW, Roesch MR (2012) Ventral striatum encodes past and predicted value independent of motor contingencies. J Neurosci 32:2027–2036 Available at: http://www.ncbi.nlm.nih.gov/pubmed/22323717 [Accessed July 25, 2012].

Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science 301:1104–1107 Available at: http://www.ncbi.nlm.nih.gov/pubmed/12934011.

Grace AA, Floresco SB, Goto Y, Lodge DJ (2007) Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. Trends Neurosci 30:220–227.

Grant S, Contoreggi C, London ED (2000) Drug abusers show impaired performance in a laboratory test of decision making. Neuropsychologia 38:1180–1187.

Graybiel A, Aosaki T, Flaherty A, Kimura M (1994) The basal ganglia and adaptive motor control. Science (80-) 265:1826–1831 Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.8091209.

Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J Neurosci 20:2369–2382.

Haber SN, Knutson B (2009) The Reward Circuit: Linking Primate Anatomy and Human Imaging. Neuropsychopharmacology 35:4–26 Available at: http://dx.doi.org/10.1038/npp.2009.129.

Halbout B, Liu AT, Ostlund SB (2016) A closer look at the effects of repeated cocaine exposure on adaptive decision-making under conditions that promote goal-directed control. Front Psychiatry 7.

Hassani OK, Cromwell HC, Schultz W (2001) Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. J Neurophysiol 85:2477–2489.

Heil SH, Johnson MW, Higgins ST, Bickel WK (2006) Delay discounting in currently using and currently abstinent cocaine-dependent outpatients and non-drug-using matched controls. Addict Behav 31:1290–1294.

Hernandez A, Burton AC, O'Donnell P, Schoenbaum G, Roesch MR (2015) Altered Basolateral Amygdala Encoding in an Animal Model of Schizophrenia. J Neurosci 35:6394–6400 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5096-14.2015.

Herrnstein RJ (1961) Relative and absolute strength of response as a function of frequency of reinforcement. J Exp Anal Behav 4:267–272.

Hikida T, Kimura K, Wada N, Funabiki K, Nakanishi Shigetada S (2010) Distinct Roles of Synaptic Transmission in Direct and Indirect Striatal Pathways to Reward and Aversive Behavior. Neuron 66:896–907.

Hikosaka K, Watanabe M (2004) Long- and short-range reward expectancy in the primate orbitofrontal cortex. Eur J Neurosci 19:1046–1054.

Hikosaka O (2007) Basal ganglia mechanisms of reward-oriented eye movement. In: Annals of the New York Academy of Sciences, pp 229–249.

Ho MY, Mobini S, Chiang TJ, Bradshaw CM, Szabadi E (1999) Theory and method in the quantitative analysis of "impulsive choice" behaviour: Implications for psychopharmacology. Psychopharmacology (Berl) 146:362–372.

Hogarth L, Balleine BW, Corbit LH, Killcross S (2013) Associative learning mechanisms underpinning the transition from recreational drug use to addiction. Ann N Y Acad Sci 1282:12–24.

Holland PC (2004) Relations Between Pavlovian-Instrumental Transfer and Reinforcer Devaluation. J Exp Psychol Anim Behav Process 30:104–117 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/0097-7403.30.2.104.

Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior-related neuronal activity in primate striatum. J Neurophysiol 80:947–963.

Houk JC, Adams JL, Barto AG (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Models of Information Processing in the Basal Ganglia, pp 249–270.

Hyman SE, Malenka RC, Nestler EJ (2006) Neural Mechanisms of Addiction : The Role of Reward-Related Learning and Memory. Annu Rev Neurosci 29:565–598.

Ikemoto S (2007) Dopamine reward circuitry: Two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. Brain Res Rev 56:27–78.

Ikemoto S, Panksepp J (1999) The role of nucleus accumbens dopamine in motivated behavior: A unifying interpretation with special reference to reward-seeking. Brain Res Rev 31:6–41.

Ito M, Doya K (2009) Validation of decision-making models and analysis of decision variables in the rat basal ganglia. J Neurosci 29:9861–9874 Available at: http://www.jneurosci.org/cgi/content/full/29/31/9861%5Cnpapers3://publication/doi/10.1523/JNEUROSCI.6157-08.2009.

Ito R, Dalley JW, Howes SR, Robbins TW, Everitt BJ (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. J Neurosci 20:7489–7495 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11007908.

Ito R, Dalley JW, Robbins TW, Everitt BJ (2002) Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. J Neurosci 22:6247–6253 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12122083.

Ito R, Robbins TW, Everitt BJ (2004) Differential control over cocaine-seeking behavior by nucleus accumbens core and shell. Nat Neurosci 7:389–397.

Izquierdo A, Suda RK, Murray EA (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J Neurosci 24:7540–7548.

Janak PH, Chen MT, Caulder T (2004) Dynamics of neural coding in the accumbens during extinction and reinstatement of rewarded behavior. Behav Brain Res 154:125–135.

Jentsch JD, Olausson P, De La Garza R, Taylor JR (2002) Impairments of reversal learning and response perseveration after repeated, intermittent cocaine administrations to monkeys. Neuropsychopharmacology 26:183–190.

Jentsch JD, Taylor JR (1999) Impulsivity resulting from frontostriatal dysfunction in drug abuse: Implications for the control of behavior by reward-related stimuli. Psychopharmacology (Berl) 146:373–390.

Jimura K, Chushak MS, Braver TS (2013) Impulsivity and Self-Control during Intertemporal Decision Making Linked to the Neural Dynamics of Reward Value Representation. J Neurosci 33:344–357 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0919-12.2013.

Joel D, Niv Y, Ruppin E (2002) Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Networks 15:535–547.

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel a M (1999) Building neural representations of habits. Science 286:1745–1749.

Jones B, Mishkin M (1972) Limbic lesions and the problem of stimulus-Reinforcement associations. Exp Neurol 36:362–377.

Kahneman D, Tversky A (1984) Choices, values, and frames. Am Psychol 39:341–350 Available at: http://psycnet.apa.org/journals/amp/39/4/341.html%5Cnpapers3://publication/doi/10.1037/0003-066X.39.4.341.

Kalenscher T, Pennartz CMA (2008) Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. Prog Neurobiol 84:284–315.

Kalenscher T, Windmann S, Diekamp B, Rose J, G??nt??rk??n O, Colombo M (2005) Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. Curr Biol 15:594–602.

Keiflin R, Janak PH (2015) Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. Neuron 88:247–263.

Kelz MB et al. (1999) Expression of the transcription factor deltaFosB in the brain controls sensitivity to cocaine. Nature 401:272–276 Available at: http://www.ncbi.nlm.nih.gov/pubmed/10499584%5Cnhttp://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10499584.

Khamassi M, Mulder AB, Tabuchi E, Douchamps V, Wiener SI (2008) Anticipatory reward signals in ventral striatal neurons of behaving rats. Eur J Neurosci 28:1849–1866.

Kim J, Ragozzino ME (2005) The involvement of the orbitofrontal cortex in learning under changing task contingencies. Neurobiol Learn Mem 83:125–133.

Koob GF, Volkow ND (2010) Neurocircuitry of addiction. Neuropsychopharmacology 35:217–238 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19710631%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2805560.

Kosaki Y, Dickinson A (2010) Choice and contingency in the development of

behavioral autonomy during instrumental conditioning. J Exp Psychol Anim Behav Process 36:334–342 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0016887.

Kreitzer AC, Malenka RC (2008) Striatal Plasticity and Basal Ganglia Circuit Function. Neuron 60:543–554.

Kupchik YM, Brown RM, Heinsbroek JA, Lobo MK, Schwartz DJ, Kalivas PW (2015) Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. Nat Neurosci 18:1230–1232 Available at: http://www.nature.com/doifinder/10.1038/nn.4068.

Lau B, Glimcher PW (2007) Action and Outcome Encoding in the Primate Caudate Nucleus. J Neurosci 27:14502–14514 Available at: http://www.jneurosci.org/content/27/52/14502%5Cnhttp://www.jneurosci.org/content/27/52/14502.full.pdf%5Cnhttp://www.jneurosci.org/content/27/52/14502.long%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/18160658.

Lau B, Glimcher PW (2008) Value Representations in the Primate Striatum during Matching Behavior. Neuron 58:451–463.

Lauwereyns J, Watanabe K, Coe B, Hikosaka O (2002) A neural correlate of response bias in monkey caudate nucleus. Nature 418:413–417.

LeBlanc KH, Maidment NT, Ostlund SB (2013) Repeated Cocaine Exposure Facilitates the Expression of Incentive Motivation and Induces Habitual Control in Rats. PLoS One 8.

Lee BR, Ma Y-Y, Huang YH, Wang X, Otaka M, Ishikawa M, Neumann PA, Graziane NM, Brown TE, Suska A, Guo C, Lobo MK, Sesack SR, Wolf ME,

Nestler EJ, Shaham Y, Schlüter OM, Dong Y (2013) Maturation of silent synapses in amygdala-accumbens projection contributes to incubation of cocaine craving. Nat Neurosci 16:1644–1651 Available at: http://www.nature.com/doifinder/10.1038/nn.3533.

Liljeholm M, O'Doherty JP (2012) contributions of the striatum to learning, motivation, and performance: An associative account. Trends Cogn Sci 16:1364–6613.

Lobo MK, Covington HE, Chaudhury D, Friedman AK, Sun H, Damez-Werno D, Dietz DM, Zaman S, Koo JW, Kennedy PJ, Mouzon E, Mogri M, Neve RL, Deisseroth K, Han M-H, Nestler EJ (2010) Cell Type-Specific Loss of BDNF Signaling Mimics Optogenetic Control of Cocaine Reward. Science (80-) 330:385–390 Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1188472.

Loewenstein G, Prelec D (1992) Anomalies in Intertemporal Choice: Evidence and an Interpretation. Q J Econ 107:573–597 Available at: http://www.jstor.org/stable/2118482.

Lucantonio F, Caprioli D, Schoenbaum G (2014) Transition from "model-based" to "model-free" behavioral control in addiction: Involvement of the orbitofrontal cortex and dorsolateral striatum. Neuropharmacology 76:407–415.

Lucantonio F, Stalnaker T a, Shaham Y, Niv Y, Schoenbaum G (2012) The impact of orbitofrontal dysfunction on cocaine addiction. Nat Neurosci 15:358–366 Available at: http://dx.doi.org/10.1038/nn.3014.

Lüscher C (2016) The Emergence of a Circuit Model for Addiction. Annu Rev

Neurosci:257–276 Available at: http://www.ncbi.nlm.nih.gov/pubmed/27145911.

Mackillop J (2013) Integrating behavioral economics and behavioral genetics: Delayed reward discounting as an endophenotype for addictive disorders. J Exp Anal Behav 99:14–31.

Maren S (2000) Auditory fear conditioning increases CS-elicited spike firing in lateral amygdala neurons even after extensive overtraining. Eur J Neurosci 12:4047–4054.

Matamales M, Bertran-Gonzalez J, Salomon L, Degos B, Deniau JM, Valjent E, Hervé D, Girault JA (2009) Striatal medium-sized spiny neurons: Identification by nuclear staining and study of neuronal subpopulations in BAC transgenic mice. PLoS One 4.

Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459:837–841.

McAlonan K, Brown VJ (2003) Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. Behav Brain Res 146:97–103.

McDannald M a, Lucantonio F, Burke K a, Niv Y, Schoenbaum G (2011) Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. J Neurosci 31:2700–2705 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3079289&tool=pmcentrez&rendertype=abstract.

Meck WH, Penney TB, Pouthas V (2008) Cortico-striatal representation of time in animals and humans. Curr Opin Neurobiol 18:145–152.

Mello GBM, Soares S, Paton JJ (2015) A scalable population code for time in the striatum. Curr Biol 25:1113–1122.

Mendez IA, Simon NW, Hart N, Mitchell MR, Nation JR, Wellman PJ, Setlow B (2010) Self-administered cocaine causes long-lasting increases in impulsive choice in a delay discounting task. Behav Neurosci 124:470–477 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2976632&tool=pmcentrez&rendertype=abstract.

Meunier M, Bachevalier J, Mishkin M (1997) Effects of orbital frontal and anterior cingulate lesions on object and spatial memory in rhesus monkeys. Neuropsychologia 35:999–1015.

Mitchell MR, Weiss VG, Ouimet DJ, Fuchs RA, Morgan D, Setlow B (2014) Intake-dependent effects of cocaine self-administration on impulsive choice in a delay discounting task. Behav Neurosci 128:419–429 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4107092&tool=pmcentrez&rendertype=abstract.

Mobini S, Body S, Ho MY, Bradshaw C, Szabadi E, Deakin J, Anderson I (2002) Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology (Berl) 160:290–298.

Mogenson GJ, Jones DL, Yim CY (1980) From motivation to action: Functional interface between the limbic system and the motor system. Prog Neurobiol 14:69–97.

Nakamura K, Santos GS, Matsuzaki R, Nakahara H (2012) Differential reward coding in the subdivisions of the primate caudate during an oculomotor task. J

Neurosci 32:15963–15982 Available at: http://www.ncbi.nlm.nih.gov/pubmed/23136434.

Nelson A, Killcross S (2006) Amphetamine Exposure Enhances Habit Formation. J Neurosci 26:3805–3812 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4305-05.2006.

Nelson AJD, Killcross S (2013) Accelerated habit formation following amphetamine exposure is reversed by D1, but enhanced by D2, receptor antagonists. Front Neurosci.

Nestler EJ (2013) Cellular basis of memory for addiction. Dialogues Clin Neurosci 15:431–443.

Nicola SM (2007) The nucleus accumbens as part of a basal ganglia action selection circuit. Psychopharmacology (Berl) 191:521–550.

Nicola SM, Taha SA, Kim SW, Fields HL (2005) Nucleus accumbens dopamine release is necessary and sufficient to promote the behavioral response to reward-predictive cues. Neuroscience 135:1025–1033.

Nishijo H, Yamamoto Y, Ono T, Uwano T, Yamashita J, Yamashima T (1997) Single neuron responses in the monkey anterior cingulate cortex during visual discrimination. Neurosci Lett 227:79–82 Available at: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6T0G-3PFB08W-21&_user=10&_coverDate=05%2F16%2F1997&_rdoc=1&_fmt=high&_orig=search&_origin=search&_sort=d&_docanchor=&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=1c57855a35fad987.

Nishizawa K, Fukabori R, Okada K, Kai N, Uchigashima M, Watanabe M, Shiota a., Ueda M, Tsutsui Y, Kobayashi K (2012) Striatal Indirect Pathway Contributes to Selection Accuracy of Learned Motor Actions. J Neurosci 32:13421–13432.

Niv Y, Schoenbaum G (2008) Dialogues on prediction errors. Trends Cogn Sci 12:265–272.

Nordquist RE, Voorn P, de Mooij-van Malsen JG, Joosten RNJMA, Pennartz CMA, Vanderschuren LJMJ (2007) Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. Eur Neuropsychopharmacol 17:532–540.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable Role of Ventral and Dorsal Striatum in Instrumental Conditioning. Science (80-) 304:452–454 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15087550.

O'Doherty JP (2004) Reward representations and reward-related learning in the human brain: Insights from neuroimaging. Curr Opin Neurobiol 14:769–776.

O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. Neuron 33:815–826.

Odum AL (2011) Delay Discounting: I'm a k, You're a k. J Exp Anal Behav 96:427–439 Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3213005/.

Ostlund SB, Balleine BW (2008) On habits and addiction: an associative analysis of compulsive drug seeking. Drug Discov Today Dis Model 5:235–245.

Padoa-Schioppa C (2011) Neurobiology of Economic Choice: A Good-Based Model. Annu Rev Neurosci 34:333–359 Available at:

http://www.annualreviews.org/doi/10.1146/annurev-neuro-061010-113648.

Parker NF, Cameron CM, Taliaferro JP, Lee J, Choi JY, Davidson TJ, Daw ND, Witten IB (2016) Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. Nat Neurosci 19:845–854 Available at: http://www.nature.com/doifinder/10.1038/nn.4287.

Phillips GD, Vugler A (2011) Effects of sensitization on the detection of an instrumental contingency. Pharmacol Biochem Behav 100:48–58.

Quirk GJ, Repa JC, LeDoux JE (1995) Fear conditioning enhances short-latency auditory responses of lateral amygdala neurons: Parallel recordings in the freely behaving rat. Neuron 15:1029–1039.

Ragozzino ME, Ragozzino KE, Mizumori SJY, Kesner RP (2002) Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. Behav Neurosci 116:105–115 Available at: http://www.ncbi.nlm.nih.gov/pubmed/11898801%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3273324.

Redish a D (2004a) Addiction as a computational process gone awry. Science 306:1944–1947.

Redish AD (2004b) Addiction as a Computational Process Gone Awry. Science (80-) 306:1944–1947 Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1102384.

Redish AD, Jensen S, Johnson A (2008) A unified framework for addiction: Vulnerabilities in the decision process. Behav Brain Sci 31:415–487 Available at: http://www.journals.cambridge.org/abstract_S0140525X0800472X.

Reynolds B (2006) A review of delay-discounting research with humans: relations to drug use and gambling. Behav Pharmacol 17:651–667 Available at: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00008877-200612000-00002.

Richards JB, Mitchell SH, de Wit H, Seiden LS (1997) Determination of discount functions in rats with an adjusting-amount procedure. J Exp Anal Behav 67:353–366.

Robbins TW, Ersche KD, Everitt BJ (2008) Drug addiction and the memory systems of the brain. Ann N Y Acad Sci 1141:1–21.

Robbins TW, Everitt BJ (1999) Drug addiction: bad habits add up. Nature 398:567–570.

Roberts AC (2006) Primate orbitofrontal cortex and adaptive behaviour. Trends Cogn Sci 10:83–90.

Robison AJ, Nestler EJ (2011) Transcriptional and epigenetic mechanisms of addiction. Nat Rev Neurosci 12:623–637 Available at: http://www.nature.com/doifinder/10.1038/nrn3111.

Rodriguez ML, Logue a W (1988) Adjusting delay to reinforcement: comparing choice in pigeons and humans. J Exp Psychol Anim Behav Process 14:105–117.

Roesch MR, Bryden DW (2011) Impact of size and delay on neural activity in the rat limbic corticostriatal system. Front Neurosci.

Roesch MR, Calu DJ, Schoenbaum G (2007a) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat Neurosci 10:1615–1624.

Roesch MR, Olson CR (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. Science 304:307–310 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15073380.

Roesch MR, Olson CR (2005) Neuronal activity in primate orbitofrontal cortex reflects the value of time. J Neurophysiol 94:2457–2471.

Roesch MR, Singh T, Brown PL, Mullins SE, Schoenbaum G (2009) Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. J Neurosci 29:13365–13376.

Roesch MR, Takahashi Y, Gugsa N, Bissonette GB, Schoenbaum G (2007b) Previous cocaine exposure makes rats hypersensitive to both delay and reward magnitude. J Neurosci 27:245–250 Available at: http://www.jneurosci.org/content/27/1/245.short.

Roesch MR, Taylor AR, Schoenbaum G (2006) Encoding of Time-Discounted Rewards in Orbitofrontal Cortex Is Independent of Value Representation. Neuron 51:509–520.

Rogers RD, Everitt BJ, Baldacchino A, Blackshaw AJ, Swainson R, Wynne K, Baker NB, Hunter J, Carthy T, Booker E, London M, Deakin JFW, Sahakian BJ, Robbins TW (1999) Dissociable deficits in the decision-making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteersEvidence for monoaminergic mechanisms. Neuropsychopharmacology 20:322–339.

Rogers RD, Robbins TW (2001) Investigating the neurocognitive deficits associated with chronic drug misuse. Curr Opin Neurobiol 11:250–257.

Roitman MF, Wheeler RA, Carelli RM (2005) Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. Neuron 45:587–597.

Rolls ET, Hornak J, Wade D, McGrath J (1994) Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J Neurol Neurosurg Psychiatry 57:1518–1524.

Saddoris MP, Carelli RM (2014) Cocaine self-administration abolishes associative neural encoding in the nucleus accumbens necessary for higher-order learning. Biol Psychiatry 75:156–164.

Saddoris MP, Gallagher M, Schoenbaum G (2005) Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. Neuron 46:321–331.

Saddoris MP, Sugam JA, Carelli RM (2016a) Prior Cocaine Experience Impairs Normal Phasic Dopamine Signals of Reward Value in Accumbens Shell. Neuropsychopharmacology:1–27 Available at: http://www.ncbi.nlm.nih.gov/pubmed/27604567.

Saddoris MP, Wang X, Sugam JA, Carelli RM (2016b) Cocaine Self-Administration Experience Induces Pathological Phasic Accumbens Dopamine Signals and Abnormal Incentive Behaviors in Drug-Abstinent Rats. J Neurosci 36:235–250 Available at: http://www.jneurosci.org/content/36/1/235.full.

Sanghera MK, Rolls ET, Roper-Hall A (1979) Visual responses of neurons in the dorsolateral amygdala of the alert monkey. Exp Neurol 63:610–626.

Schmitzer-Torbert N, Apostolidis S, Amoa R, O'Rear C, Kaster M, Stowers J, Ritz R

(2015) Post-training cocaine administration facilitates habit learning and requires the infralimbic cortex and dorsolateral striatum. Neurobiol Learn Mem 118:105–112.

Schmitzer-Torbert N, Redish AD (2004) Neuronal Activity in the Rodent Dorsal Striatum in Sequential Navigation: Separation of Spatial and …. J Neurophysiol Available at: http://jn.physiology.org/cgi/content/full/91/5/2259%5Cnpapers://27662af5-e24b-497a-8780-4dcec04ea037/Paper/p2313.

Schoenbaum G, Chiba a a, Gallagher M (1999) Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. J Neurosci 19:1876–1884.

Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nat Neurosci 1:155–159 Available at: http://www.nature.com/doifinder/10.1038/407%5Cnpapers3://publication/doi/10.1038/407.

Schoenbaum G, Saddoris MR, Ramus SJ, Shaham Y, Setlow B (2004) Cocaine-experienced rats exhibit learning deficits in a task sensitive to orbitofrontal cortex lesions. Eur J Neurosci 19:1997–2002.

Schoenbaum G, Setlow B (2005) Cocaine makes actions insensitive to outcomes but not extinction: Implications for altered orbitofrontal-amygdalar function. Cereb Cortex 15:1162–1169.

Schoenbaum G, Setlow B, Nugent S, Saddoris M, Gallagher M (2003a) Lesions of

orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn Mem 10:129–140 Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-0037360529&partnerID=40&md5=4ba7612c2675d73aa071c96b65257ec4.

Schoenbaum G, Setlow B, Saddoris MP, Gallagher M (2003b) Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. Neuron 39:855–867.

Schultz W (2006) Behavioral theories and the neurophysiology of reward. Annu Rev Psychol 57:87–115 Available at: http://www.ncbi.nlm.nih.gov/pubmed/16318590.

Schultz W (2013) Updating dopamine reward signals. Curr Opin Neurobiol 23:229–238.

Setlow B, Mendez IA, Mitchell MR, Simon NW (2009) Effects of chronic administration of drugs of abuse on impulsive choice (delay discounting) in animal models. Behav Pharmacol 20:380–389 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2874684&tool=pmcentrez&rendertype=abstract.

Setlow B, Schoenbaum G, Gallagher M (2003) Neural encoding in ventral striatum during olfactory discrimination learning. Neuron 38:625–636 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12765613.

Simon NW, Mendez IA, Setlow B (2007) Cocaine exposure causes long-term increases in impulsive choice. Behav Neurosci 121:543–549 Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2581406&tool=pmcentrez&rendertype=abstract.

Singh T, McDannald M a, Takahashi YK, Haney RZ, Cooch NK, Lucantonio F, Schoenbaum G (2011) The role of the nucleus accumbens in knowing when to respond. Learn Mem 18:85–87.

Singh T, McDannald MA, Haney RZ, Cerri DH, Schoenbaum G (2010) Nucleus Accumbens Core and Shell are Necessary for Reinforcer Devaluation Effects on Pavlovian Conditioned Responding. Front Integr Neurosci 4:1–7 Available at: http://www.ncbi.nlm.nih.gov/pubmed/21088698.

Son J-H, Latimer C, Keefe KA (2011) Impaired Formation of Stimulus–Response, But Not Action–Outcome, Associations in Rats with Methamphetamine-Induced Neurotoxicity. Neuropsychopharmacology 36:2441–2451.

Stalnaker T a, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G (2010) Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Front Integr Neurosci 4:12.

Stalnaker TA, Franz TM, Singh T, Schoenbaum G (2007a) Basolateral Amygdala Lesions Abolish Orbitofrontal-Dependent Reversal Impairments. Neuron 54:51–58.

Stalnaker TA, Roesch MR, Franz TM, Burke KA, Schoenbaum G (2006) Abnormal associative encoding in orbitofrontal neurons in cocaine-experienced rats during decision-making. Eur J Neurosci 24:2643–2653.

Stalnaker TA, Roesch MR, Franz TM, Calu DJ, Singh T, Schoenbaum G (2007b) Cocaine-induced decision-making deficits are mediated by miscoding in

basolateral amygdala. Nat Neurosci 10:949–951 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=17603478 http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17603478%5Cnhttp://www.nature.com/neuro/journal/v10/n8/pdf.

Stalnaker TA, Takahashi Y, Roesch MR, Schoenbaum G (2009) Neural substrates of cognitive inflexibility after chronic cocaine exposure. Neuropharmacology 56:63–72.

Steinberg EE, Boivin JR, Saunders BT, Witten IB, Deisseroth K, Janak PH (2014) Positive reinforcement mediated by midbrain dopamine neurons requires D1 and D2 receptor activation in the nucleus accumbens. PLoS One 9.

Sutton RS, Barto AG (1998) Reinforcement Learning : An Introduction.

Tai L-H, Lee AM, Benavidez N, Bonci A, Wilbrecht L (2012) Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. Nat Neurosci 15:1281–1289 Available at: http://www.nature.com/doifinder/10.1038/nn.3188.

Takahashi Y, Roesch MR, Stalnaker T a, Schoenbaum G (2007) Cocaine exposure shifts the balance of associative encoding from ventral to dorsolateral striatum. Front Integr Neurosci 1:11.

Takahashi Y, Schoenbaum G, Niv Y (2009) Silencing the critics: Understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an Actor/Critic model. Front Neurosci 2:86–89.

Takahashi YK, Langdon AJ, Niv Y, Schoenbaum G (2016) Temporal Specificity of Reward Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum. Neuron 91:182–193.

Takahashi YK, Schoenbaum G (2016) Ventral striatal lesions disrupt dopamine neuron signaling of differences in cue value caused by changes in reward timing but not number. Behav Neurosci 130:593–599 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/bne0000169.

Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2016) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. In: Behavioral Economics of Preferences, Choices, and Happiness, pp 593–616.

Thaler RH (1981) Some Empirical Evidence on Dynamic Inconsistency. Econ Lett 8:201–207 Available at: http://linkinghub.elsevier.com/retrieve/pii/0165176581900677.

Thomsen M, Caine SB (2005) Chronic intravenous drug self-administration in rats and mice. Curr Protoc Neurosci Chapter 9:Unit 9.20.

Toyomitsu Y, Nishijo H, Uwano T, Kuratsu J, Ono T (2002) Neuronal responses of the rat amygdala during extinction and reassociation learning in elementary and configural associative tasks. Eur J Neurosci 15:753–768.

Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature 398:704–708 Available at: http://www.nature.com/doifinder/10.1038/19525%5Cnpapers3://publication/doi/10.1038/19525.

Tricomi E, Balleine BW, O'Doherty JP (2009) A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci 29:2225–2232.

Ungless MA, Whistler JL, Malenka RC, Bonci A (2001) Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. Nature 411:583–7.

van der Meer MAA, Johnson A, Schmitzer-Torbert NC, Redish AD (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 MAA, Redish AD (2011) Ventral striatum: A critical look at models of learning and evaluation. Curr Opin Neurobiol 21:387–392.

Van der Meer MA, Redish AD (2009) Covert Expectation-of-Reward in Rat Ventral Striatum at Decision Points. Front Integr Neurosci 3:1–15.

Vanderschuren LJ, Di Ciano P, Everitt BJ (2005) Involvement of the dorsal striatum in cue-controlled cocaine seeking. J Neurosci 25:8665–8670 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16177034.

Volkow ND, Baler RD (2015) NOW vs LATER brain circuits: Implications for obesity and addiction. Trends Neurosci 38:345–352.

Volkow ND, Fowler JS, Wang G-J, Swanson JM (2004) Dopamine in drug abuse and addiction: results from imaging studies and treatment implications. Mol Psychiatry 9:557–569 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15098002.

Volkow ND, Morales M (2015) The Brain on Drugs: From Reward to Addiction. Cell 162:712–725.

Volkow ND, Wang G-J, Fowler JS, Tomasi D (2015) Addiction Circuitry in the Human Brain. Focus (Madison) 13:341–350 Available at: http://psychiatryonline.org/doi/10.1176/appi.focus.130306.

Wied HM, Jones JL, Cooch NK, Berg BA, Schoenbaum G (2013) Disruption of model-based behavior and learning by cocaine self-administration in rats. Psychopharmacology (Berl) 229:493–501.

Willuhn I, Burgeno LM, Everitt BJ, Phillips PEM (2012) Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. Proc Natl Acad Sci U S A 109:20703–20708 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3528544&tool=pmcentrez&rendertype=abstract.

Winstanley CA (2004) Contrasting Roles of Basolateral Amygdala and Orbitofrontal Cortex in Impulsive Choice. J Neurosci 24:4718–4722 Available at: http://www.jneurosci.org.ezproxy.otago.ac.nz/content/24/20/4718.

Wise RA, Koob GF (2014) The development and maintenance of drug addiction. Neuropsychopharmacology 39:254–262 Available at: http://www.nature.com/npp/journal/v39/n2/full/npp2013261a.html#bib67.

Wolf ME, Ferrario CR (2010) AMPA receptor plasticity in the nucleus accumbens after repeated exposure to cocaine. Neurosci Biobehav Rev 35:185–211.

Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nat Rev Neurosci 7:464–476 Available at: http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&DbFrom=pubmed&Cmd=Link&LinkName=pubmed_pubmed&LinkReadableName=Related

Articles&IdsFromResult=16715055&ordinalpos=3&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum.

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181–189.

Zachariou V, Bolanos CA, Selley DE, Theobald D, Cassidy MP, Kelz MB, Shaw-Lutchman T, Berton O, Sim-Selley LJ, Dileone RJ, Kumar A, Nestler EJ (2006) An essential role for DeltaFosB in the nucleus accumbens in morphine action. Nat Neurosci 9:205–211 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16415864.