Decision Making Gone Awry:
Dorsal striatum, decision-making, and addiction

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Paul Regier

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

A. David Redish

February 2015

© Paul S. Regier 2015

Acknowledgements

I would like to acknowledge everyone that helped to make this research and dissertation possible. Thank you to everyone in Dr. David Redish's lab, especially Dr. Redish for his support and expertise, guiding me during the last three years of my Ph.D, and making me a better scientist. Thank you to everyone in Dr. Robert Meisel's lab, who helped me with my very first project in graduate school and, subsequently, my first publication. Thank you to everyone in Dr. Marilyn Carroll's lab, especially Dr. Anker, Dr. Holtz, and Dr. Zlebnik. Thank you to my committee members for challenging me and making my dissertation even better. Finally, thank you to all of my friends and family members, who have been ever loving and supportive.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Addiction is similar to other behaviors, driven by the same mechanisms that assess value, motivate, and elicit want and desire. Even though there are different factors that have been thought to contribute to the risk of an individual engaging in addictive behaviors (Wise, 1996; Koob et al., 1998; Nestler, 2001; Robinson and Berridge, 2001), there appear to be common underlying brain correlates that influence the decision-making process as it relates to addiction (Redish et al., 2008). However, in order to understand decision-making that is maladaptive, such as addictive behavior, we need to first understand normal decision-making.

Both topics of decision making and addiction are extensive and have been reviewed elsewhere (Barto, 1998; Koob et al., 1998; Everitt and Robbins, 2005; Dayan et al., 2006; Niv et al., 2006; Redish et al., 2008; van der Meer et al., 2012; Everitt and Robbins, 2013; Redish, 2013). In my studies, although I employed several different research techniques (e.g. behavioral pharmacology, immunohistochemistry, neurophysiology, and theoretical neuroscience), it was my purpose to focus on and relate findings to one brain area, the dorsal striatum and subsections of the dorsal striatum.

Subsections of the dorsal striatum have been found to be related to flexible and habit-based decision-making, as well as decision making related to addiction. This is especially true when considering animal models of addiction that differentially seek drugs. Reports from studies in the dorsal striatum as they related to normal and abnormal

decision-making have important implications for the dorsal striatum as a biomarker for treatment of addiction as well as translational relevance for treatment of human addiction.

In the following dissertation, I will, first, introduce the dorsal striatum and decision-making systems and discuss the importance of the dorsal striatum in decision-making, in general. In the next chapter, I will present a new study illuminating how both dorsomedial and dorsolateral striatum are correlated with flexible and habit-based decision-making processes, respectively, during navigation on a spatial task. Then, in chapter 3, I will discuss several theories of addiction and how dorsal striatum is involved with these theories. In chapter 4, I will discuss specific animal models of addiction, how animal models that self-administer more or less drugs differ neurobiologically, as well as respond differently to punishment paired with drugs. I will end the chapter with a discussion about how differential dorsostriatal functioning might influence divergent drug-seeking behavior, especially those behaviors that continue even in spite of negative consequences. Chapter 5 will address how these animal models differ in treatment receptivity, with a discussion of how differential dorsostriatal functioning might influence treatment receptivity. In chapter 6, I will relate results from previous studies and results of my studies to human drug use and how analogues of dorsal striatum in the human brain might play a role in human addiction and addiction treatment. The chapter will focus on a successful treatment for addiction, Contingency Management, as an example of utilizing different decision-making systems in order to increase treatment receptivity. Finally, in chapter 7, I will summarize findings, showing a clear path from neurobiology to human behavior as it relates to addiction.

Dorsal Striatum

The dorsal striatum is a part of the basal ganglia, which also includes the globus pallidus, substantia nigra, thalamus, and cortex. Early studies with non-human primates discovered discrete, essentially, non-overlapping basal ganglia-thalamocortical circuits (Alexander et al., 1986; Alexander and Crutcher, 1990): the motor circuit, the oculomotor circuit, two prefrontal circuits, and the limbic circuit. Research found that multiple inputs from cortex to caudate/putamen were progressively integrated, forming partially closed circuit loops, until the loops eventually terminated back on the cortex (Alexander et al., 1986). Research found that neurons in the caudate/putamen were grouped together in clusters, or striosomes (Graybiel et al., 1981), and, at least in the putamen, when stimulated, activated specific body parts. Output from the striatum was long thought to form two distinct pathways: the direct (movement-releasing) and indirect (movement inhibiting) pathways. However, recent data has suggested that heavy collateralization in non-human primate caudate/putamen and rodent striatum is making concepts of striatal circuitry more complex (for review, see Graybiel, 2005)

In non-human primates there is a clear boundary, the internal capsule, between caudate and putamen, both of these areas associated with different processes. The putamen is essentially a part of the motor circuit (Alexander et al., 1986), it regulates motor movement (Alexander and Crutcher, 1990), and it is important for habit formation (Graybiel, 1998). The caudate is essentially a part of the prefrontal circuits (Alexander et al., 1986) and is involved with flexible decision-making (Devan et al., 1999; Yin and Knowlton, 2004; Yin et al., 2005; Thorn et al., 2010; Chapter 2).

In rats, there is no clear dividing line in the dorsal striatum. Even so, the dorsal striatum does not form a homogenous region. Recent studies have discovered anatomical and functional differences between dorsolateral striatum (analog of the monkey dorsal putamen) and dorsomedial striatum (analog of the monkey dorsal caudate) as well as anatomical and functional differences between *anterior* dorsomedial striatum and *posterior* dorsomedial striatum. The dorsolateral striatum receives input primarily from motor and sensory cortical areas (Alexander et al., 1986; Parent, 1986; McGeorge and Faull, 1989; Berendse et al., 1992; Swanson, 2000). Dorsolateral functioning is similar to the putamen and regulates motor control and habit-based behaviors (Alexander et al., 1986; Graybiel, 2008). The dorsomedial striatum receives projections from different pre-frontal and orbitofrontal cortical areas depending on the anterior to posterior position of the region. For example, anterior dorsomedial striatum receives input from the ventral and dorsal pre-limbic areas and some input from ventral orbitofrontal cortex (McGeorge and Faull, 1989; Berendse et al., 1992; Swanson, 2000). The posterior dorsomedial striatum receives input from ventral orbitofrontal cortex and the dorsal pre-limbic area and some input from the ventral pre-limbic area and the infralimbic area (McGeorge and Faull, 1989; Berendse et al., 1992; Swanson, 2000). Dorsomedial functioning is similar to the monkey caudate, but differs depending on posterior or anterior orientation. Detailed discussion of these differences appears in the next chapter.

Decision-making Systems

When an agent encounters a new environment, the agent naturally learns associations between stimuli in the environment and potential rewards. In order to do this, an agent

needs to be able to process information on how to obtain and maximize reward intake. In this dissertation I will focus on two main decision-making processes[1]: goal-directed behavior and habit-based behavior. These are two distinct decision-making processes that process information on how to assess value and obtain rewards in very different ways. Goal-directed behavior evaluates action and its outcome, processing information about past and potential rewards, at the time of action selection, whereas habit-based behavior evaluates the value of each action, processing information about past and present received rewards, at the time of reward receipt.

Studies with rats have shown that early in behavioral training, rats engage in goal-directed behavior, wherein actions made are contingent upon the outcome, but after extensive training on a task, the rat's actions are no longer contingent upon the outcome

[1] Because of the focus on dorsal striatum, Pavlovian decision-making will only be mentioned, at times, but not fully explored. This is due to dorsal striatum being implicated as playing a major role in habit-based and goal-directed decision-making but not Pavlovian decision-making. Pavlovian decision-making involves hard-wired species-specific actions that are released by associated stimuli (van der Meer et al., 2012a). Ventral striatum, amygdala, ventral tegmental area, and orbitofrontal cortex have been found to underlie Pavlovian decision-making (Cardinal et al., 2002). It has been also been related to theories of addiction (Chapter 3), such as noncompensable dopamine and incentive salience (Berridge and Robinson, 1998; Redish et al., 2008; van der Meer et al., 2012). Furthermore, maladaptive Pavlovian conditioning has been found to influence increased drug-seeking behaviors in rodent models of addiction (Chapters 4 and 5), predicted by the association of more value to the cues that predict drug reward rather than the rewards themselves (Flagel et al., 2011). Thus, a cue associated with an outcome (e.g., drug consumption) releases a behavior (e.g., approach drug cue), increasing the probability of drug use.

and have become habit-based (Packard and McGaugh, 1996; Yin and Knowlton, 2004; Niv et al., 2006).

This had been discovered, in part, by measuring how animals react to revaluation of the reward at various levels of training experience. One method of revaluation involves devaluing a food reward by lithium chloride or by satiating the animal on the food reward. For example, pairing an aversive stimulus with food decreased actions to obtain reward early in training when rats were engaged in goal-directed behavior but not after extended training, when rats were engaged in habit-based behavior (Adams and Dickinson, 1981; Adams, 1982).

Goal-directed behavior is used in early learning, since it takes into account the current state of the agent, integrating past experience with potential future experiences. The agent can make predictions about what actions might correspond with future goals. Nonetheless, because the agent engages in evaluative and predictive steps, such as which action to take and the outcome of each action, at the time of action selection, the process of goal-directed behavior is cognitively intensive. However, it is flexible, because past and potential future outcomes are considered in the process of deciding on the choice at hand (Killcross and Coutureau, 2003; Balleine et al., 2007; Buckner and Carroll, 2007; van der Meer et al., 2012b; Redish, 2013). In contrast, habit-based behavior is inflexible and relies on changes to reward value or availability of reward at the time of reward receipt in order to update information about the environment (Niv et al., 2006; Balleine et al., 2007; van der Meer et al., 2012b, 2012b). Because future outcome is not considered at the time of action selection, stimuli release actions very quickly. However, once

stimulus-action associations are well established, they are difficult to change (e.g. insensitivity to devaluation, see above).

Goal-directed and habit-based behaviors have been hypothesized to be driven by distinct but parallel-functioning, neurobiological systems (Packard and McGaugh, 1996; Yin and Knowlton, 2004; Niv et al., 2006; Balleine et al., 2007; van der Meer et al., 2012), each competing for eventual action-selection. Goal-directed behavior has been shown to have several underlying neural correlates, including the pre-frontal and orbitofrontal cortical areas, the hippocampus, and the dorsomedial striatum (Packard and McGaugh, 1996; Killcross and Coutureau, 2003; Yin and Knowlton, 2004; Balleine, 2005; Buckner and Carroll, 2007; van der Meer et al., 2012; Redish, 2013; Chapter 2). Habit-based behavior has been shown to have underlying neural correlates primarily in the motor basal ganglia-thalamocortical circuit (for reviews, see (Alexander et al., 1986; Graybiel, 2005; Graybiel, 2008)), including the motor/sensory cortices and dorsolateral striatum (Hikosaka et al., 2002; Miyachi et al., 2002; Schmitzer-Torbert and Redish, 2004; Yin and Knowlton, 2004; Balleine, 2005; Barnes et al., 2005; Yin et al., 2005; Schmitzer-Torbert and Redish, 2008; Thorn et al., 2010; Smith and Graybiel, 2013; Chapter 2).

In the next chapter, I will discuss specific studies, wherein dorsolateral and dorsomedial striatum relate to habit-based and goal-directed behavior, respectively. In addition, I will present new neurophysiological data for these two areas in relation to habit-based and goal-directed behavior in rats as they navigate a spatial navigation task.

Chapter 2. Dorsal Striatum and Decision-making Systems

The dorsal striatum is an essential part of the decision-making process. Although early work effectively treated the dorsal striatum as a whole, several studies have shown that the dorsal striatum can be functionally and anatomically separated into two distinct regions, the dorsomedial (DMS) and the dorsolateral (DLS) striatum (Devan et al., 1999; Yin and Knowlton, 2004; Yin et al., 2005a, 2006; Kimchi and Laubach, 2009; Stalnaker, 2010; Thorn et al., 2010; Stalnaker et al., 2012).

As discussed in the first chapter, both of these areas are a part of distinct basal ganglia-thalamocortical circuits, comprised of cortex, striatum, globus pallidus, substantia nigra, and thalamus (Alexander et al., 1986). While both of these regions receive some input from similar areas, such as the substantia nigra and the ventral tegmental area, the DLS receives inputs primarily from the motor and sensory cortex areas while the DMS receives input primarily from the orbitofrontal cortex, the pre-frontal cortex area, and the hippocampus (Alexander et al., 1986; Parent, 1986; McGeorge and Faull, 1989; Berendse et al., 1992; Parent and Hazrati, 1995).

Lesion studies have shown that inactivation of the DMS impairs flexible spatial navigation on a simple T-maze (Yin and Knowlton, 2004) and the Morris Water Maze (Devan et al., 1999), and the inactivation of the DLS impairs habit formation and causes subjects to revert back to flexible strategies after overtraining on a task (Packard and McGaugh, 1996; Yin and Knowlton, 2004). Yin and Knowlton (2004) also found that the DMS could be separated into distinct subregions. Inactivation of the posterior DMS (pDMS), but not anterior DMS (aDMS), disrupted the use of flexible decision making,

and they postulated that pDMS receives more input than aDMS from areas important for state recognition and learning, namely the orbitofrontal cortex and the hippocampus, respectively. Subsequent research has strengthened the differentiation of aDMS and pDMS, showing that pDMS, but not aDMS (Yin et al., 2005b) or aDLS (Yin et al., 2005a), is important for action-outcome. Although both the aDMS and pDMS receive input from the anterior cingulate cortex and the pre-limbic, the pDMS receives input from distinct areas, such as the orbitofrontal cortex and the hippocampus (McGeorge and Faull, 1989; Berendse et al., 1992; Swanson, 2000).

Research investigating the effect of lesions in the aDMS has shown that this striatal subregion is important for reversal learning (Kirkby, 1969; Kolb, 1977; Ragozzino et al., 2002a; Ragozzino and Choi, 2004; Clarke et al., 2008; Castañé et al., 2010a) and strategy switching (Ragozzino et al., 2002b; Ragozzino, 2007). The effect of inactivating the aDMS was similar to deficits found after inactivating the orbitofrontal cortex and pre-limbic area (Ferry et al., 2000; Schoenbaum et al., 2002; Boulougouris et al., 2007; Ragozzino, 2007). However, when researchers investigated the neuronal underpinnings in the aDMS of reversal learning, they found no neural correlates of reversal learning (Kimchi and Laubach, 2009).

The lack of reversal learning correlates in the aDMS is perplexing, since lesions to this region have produced reliable deficits of reversal learning. It is possible that lesions to the aDMS were not selective enough and resulted in ablation of some or all of the pDMS. Although Ragozzino et al. (2002a, 2002b, 2004) injection sites were sufficiently anterior, lesion sites from other studies show that, at least, some of the pDMS

was affected (Kirkby, 1969; Castañé et al., 2010). Prior to the seminal work of Yin and Knowlton (2004), it was not well known the rodent DMS had distinct functional subregions, and researchers may have inadvertently failed to selectively target the aDMS.

Several studies have investigated the neurophysiology of the anterior dorsal striatum on basic (Jog, 1999; Barnes et al., 2005; Berke et al., 2009; Thorn et al., 2010; Smith and Graybiel, 2013) and complex navigation tasks (Schmitzer-Torbert and Redish, 2004, 2008). Consistently, aDLS has been found to underlie more habit-based actions, developing firing patterns that underlie habit-based behaviors (more aDLS activity at the beginning and end of the maze) over several training sessions (Jog, 1999; Barnes et al., 2005; Thorn et al., 2010; Smith and Graybiel, 2013). In addition, aDLS neurons represented space more accurately with increased experience on a task within session (van der Meer et al., 2010). Development of firing rate at the beginning and end of the maze in the aDLS over sessions as the rats become more experienced on the task supported the theory of "chunking" in the aDLS. This means that aDLS activation was only required to start and end a series of actions, allocating neural control in the middle of these series of actions to other brain areas, so that corticostriatal circuits were free to learn new associations (Miller, 1956; Graybiel, 1998). Research has implicated that the aDMS may be a part of one of these corticostriatal learning circuits, since aDMS firing rate was found to be higher in the middle of a run on the maze but not at the beginning and end of trials on the maze (Thorn et al., 2010).

Even though the literature suggested that pDMS plays a role in the learning of new behaviors and strategies, neurophysiological data from pDMS has been limited in

reporting how or even if pDMS neurons might process information differently from the aDLS. Thus, it would be important to investigate information processing in both the aDLS and pDMS simultaneously, since pDMS has been implicated as playing a role in flexible decision-making behavior and the aDLS in habit-based decision-making, and compare differences.

In order to address the above-unanswered questions, we trained rats on a Hebb-Williams maze (Figure 1). Our task required rats to use spatial information and make internal-cue guided decisions at a key choice point in order to earn food rewards. During training sessions, a contingency to obtain reward was set for the entire session, where rats learned to go left, right, or alternate for food rewards. After training, rats entered the experimental phase. During six test sessions, the contingency for reward was switched mid-session, serving as a novel event and forcing rats to change their behavioral strategy. We simultaneously recorded from aDLS and pDMS and investigated neuronal response and spatial information processing across laps at specific points on the maze. We also investigated the neural correlates of a behavioral strategy change.

Study 1: Dorsal striatal subregions differently process information on a spatial navigation task

Methods for Study 1

Five Fischer Brown-Norway rats and 1 Brown-Norway Rat were trained to perform a modified version of a Hebb-Williams maze (HWM, Hebb and Williams, 1946; Rabinovitch and Rosvold, 1951), similar to the multiple-T left, right, alternate (MT-LRA) task (Blumenthal et al., 2011; Steiner and Redish, 2012; Powell and Redish, 2014). The

maze was a wooden rectangle box with carpeted floor and LEGO© walls that could be altered to change the internal maze portion, consisting of low-cost choice points, and labeled the navigation sequence (N). At the end of the navigation sequence, rats came to a high-cost choice point (C) and had to make a left or right turn. If they made the correct choice, they would receive a food reward (two unflavored food pellets, 45 mg each, Research Diets, New Brunswick, NJ, USA) at a side feeder (F) location and at a center feeder location (End Zone (E)). This started a new lap, where rats would repeat the sequence of events. If they made an incorrect choice, they did not receive any food rewards and had to continue down the return arm (R) and start over the sequence of events (Figure 1). The pellets were delivered using automatic pellet dispensers (Med-Associates, St. Albans, VT, USA). There were three different contingencies (left (L), right (R), or alternate (A)). During training sessions, the contingency was the same for the entire session. During the experimental phase, rats completed six sessions, wherein each session began with one contingency, but this contingency changed at approximately the halfway point of the session (contingency switch). Each rat ran all the possible combinations (LR, LA, RL, RA, AL, AR), and every session lasted 30 minutes. Rats earned their daily food intake on the task (~12g/day).

Surgery

After pre-training on the HWM, rats were chronically implanted with multi-tetrode hyperdrives (3 rats were implanted with 14-tetrode hyperdrives (made in house, 12 electrodes for recording, two for references), 3 rats were implanted with 28-tetrode hyperdrives (made in house, 24 electrodes for recording, 4 for references)). Three rats

were initially anesthetized with Nembutal (sodium pentobarbital, 40-50 mg/kg, Abbott Laboratories, North Chicago, IL, USA) and 3 rats were anesthetized with isoflurane. All rats were maintained on isoflurane (0.5-2% isoflurane vaporized in medical grade O_2) during the implantation. Rats were situated on a stereotaxic apparatus (Kopf) and received Dualcillin (Phoenix Pharmaceutical Inc., St. Joseph, MI, USA) intramuscularly in each hind limb. The dorsal part of the rats' heads were shaved and disinfected with alcohol (70% isopropyl) and Betadine (Purdue Rederick, Norwalk, CT, USA), and the skin overlying the scalp was removed. Several jewelers' screws were used to anchor the hyperdrive to the skull, and one of the screws was used as a recording ground. In 3 rats, 2 craniotomies (unilateral implantation of aDLS and pDMS) were opened, and in 3 rats, 4 craniotomies (bilateral implantation of aDLS and pDMS) were opened using a surgical trephine. The bundles for aDLS were centered at +0.7mm anterior of bregma and +3.5mm lateral of midline, and bundles for pDMS were centered at -0.4mm posterior of bregma and +2.5mm lateral of midline, in accordance with the study by Yin and Knowlton (2004). The craniotomies around the hyperdrive were protected with Silastic (Dow Corning, Midland, MI, USA). Dental acrylic (Perm Reline and Repair Resin, The Hygenic Corporation, Akron, OH, USA) secured the hyperdrive to the skull. Immediately after surgery, all tetrodes were turned down 640 μm. After tetrodes were turned down, rats were given subcutaneous injections (5-10 ml) of sterile saline and oral administration of Tylenol (1 ml). To prevent infections, rats received subcutaneous injections of baytril (enrofloxacin, 1.1 mg/kg) the day of surgery and for 7 days after surgery.

All procedures were conducted in accordance with National Institutes of Health guidelines for animal care and approved by IACUC at the University of Minnesota. Care was taken to minimize the number of animals used in these experiments and to minimize suffering.

Data Collection

Following surgery, tetrodes were advanced 320-640 um per day until reaching striatum. Striatum was differentiated by observation of the corpus callosum, an area that is electrophysiologically quiet compared to the cortex and striatum. The striatum is further identified by the observation of medium spiny neurons, which have long inter-spike intervals and short bursts of firing.

In 3 rats, recording neural activity while running a task was made possible by a 64 channel analog Cheetah system (Neurolynx, Tuscon, AZ, USA), and the other 3 rats a 96 channel digital Cheetah system was used. Spike trains were identified and recorded online using built-in filters, and then clustering of spike trains occurred offline. Automatic pre-clustering was accomplished using KlustaKwik (K.D. Harris), which was a list of the times at which action potentials occurred for any putative neuron, and neurons were separated into putative cells on the basis of specific waveform properties using MClust 3.5 and MClust 4.0 (A. D. Redish). To ensure the quality of each neuron, L_{ratio} and Isolation Distance (Schmitzer-Torbert et al., 2005), measures of cluster quality, were used. If clusters had L_{ratio} greater than 0.10 or Isolation Distance less than 20, the putative neurons were not used for analysis.

The position of the rat was monitored using LEDs on the head stage during

experimental recording sessions, captured by an overhead camera. Position of the rat was recorded using a video input to the Cheetah recording system, time stamping the sampled position of the LEDs. Control of the experiment was performed with Matlab. Events (e.g. feeder click and food delivery) were recorded and time stamped by the Cheetah recording system and by Matlab.

Histology

After the experiment was completed, tetrode locations were marked with small lesions by passing a small amount of anodal current (5 μA for 10s) through each tetrode. After at least two days had passed, rats were anesthetized and perfused transcardially with saline followed by 10% formalin. Brains were stored in formalin followed by 30% sucrose formalin until slicing. Coronal slices were made through the area of the implantation and stained with Cresyl Violet to visualize tetrode tracks.

Data Analysis

The proportion of time spent in long (> 2 s) interspike-intervals (ISIs) was calculated for each spike train to separate phasic from non-phasic neurons. Those ISIs were summed and divided by the total session time (Schmitzer-Torbert, 2004; Schmitzer-Torbert and Redish, 2008), providing a measure of what proportion of the session was spent in ISIs equal to or longer than 2 s. All analyses were carried out using Matlab 2012a (Mathworks).

Cell-Type Classification

Each phasically-firing neuron (PFN) was classified as being reward responsive or not or

being maze responsive or not following standard practice (Alexander et al., 1986; Kimura, 1990; Schmitzer-Torbert, 2004; Berke et al., 2009). If, in the 5 s post-feeder arrival, the mean firing rate of the PFN significantly exceeded its mean firing rate on the entire maze, it was classified as a reward neuron. To test for maze responsiveness, six points on the maze were identified (start of maze, middle of maze (navigation sequence), choice point, side feeder, return arm, and end zone), and a time window of ± 1 s was taken around each point. If the mean firing rate of PFNs at any one of the six designated points on the maze significantly exceeded its mean firing rate on the entire maze, the neuron was classified as a maze neuron.

Firing Rate

Firing rate at specific points on the maze over laps was obtained similarly to Thorn et al. (2010). Eight events on the maze were identified (start of navigation sequence, middle navigation sequence, choice point, feeder click, side feeder (enter and exit), return arm, center feeder click, and end zone (start/end of each lap)). A 2 s time window (± 1 s around each event) was used to capture firing and to create a 2-D matrix of firing rate over laps. Firing rates were z-scored by taking the mean firing rate (F) of each bin and then subtracting the mean firing rate for the rest of the maze divided by the standard deviation of firing for the rest of the maze. Z-scored firing rate was then divided into 500 ms time bins (4 bins for each event), within session over 5-lap bins in both aDLS and pDMS before and after the contingency switch. A measure of overall firing rate across laps was obtained by taking the mean z-scored firing rate for each 5-lap bin, and a measure of overall firing rate across the maze was obtained by taking the mean z-scored

firing rate for each 500 ms time bin.

Task-Bracketing Index

A measure of task-bracket-like effects was calculated based on an analysis by Smith and Graybiel (2013), who took the mean firing rate at the start and end of maze minus the mean firing rate at the auditory cue at the choice point, a measure they called the task-bracket index. Similarly, in the present study, a normalized task-bracketing index was calculated by taking the mean firing rate of the last two bins of the end zone epoch (which marked the end and beginning of each lap), then subtracting the mean firing rate from the rest of the maze and dividing by the standard deviation of the mean firing rate from the rest of the maze (z-scored the same as presented above). This was done for early laps (1-15) and late laps (16-30) before and after the switch. A 2-way ANOVA was performed to test for main effects and interactions, and post-hoc tests were performed when there was a significant main effect or interaction and corrected with Bonferroni.

Tuning Curves

A measure of tuning curves was obtained by taking neuronal firing at each time point and position of the rat for each neuron. Tuning curve information was normalized by occupancy to adjust for the amount of time the rat was at each position/time point. The maze was linearized (with 1000 points) by creating an ideal path around the maze and then associating tuning curves with those points around the maze. This was done on a lap-by-lap basis.

Correlation of Tuning Curves

After obtaining tuning curve information for aDLS and pDMS neurons on individual laps four distributions of correlation coefficients were obtained for both aDLS and pDMS regions. The four different distributions (for each condition) were obtained by taking an average of left laps before the switch vs. an average of left laps after the switch, an average of right laps before the switch vs. average of right laps after the switch, average of left vs. average of right laps before the switch, and average of left vs. average of right laps after the switch. Averages were calculated across laps, so that maze locations were preserved.

To determine whether there was a significant difference between aDLS and pDMS neuronal distributions in each of the four conditions, we conducted a bootstrap analysis. First, we randomly assigned correlation coefficients from each region for each of the four conditions. Then, correlation coefficients were selected randomly (with replacement) to create a sample distribution, and the average of each condition was computed. From this randomly sampled data, the mean distances of the x (correlation of firing rate of before vs. after switch laps) y (correlation of firing rate for left vs. right laps) coordinates from the randomly assigned aDLS to the x y coordinates of the randomly assigned pDMS were computed, providing a randomly sampled distribution of distances for each of the four conditions. We then computed the actual distance from the mean aDLS x y coordinate to the mean pDMS x y coordinate in each of the four conditions. Finally, we compared each actual distance for each condition to the randomly

sampled distribution created for each condition, and a p value was obtained by calculating the probability of random samples falling outside of actual mean distance.

Results

Behavior

Behavioral results in the present study were similar to previous studies on the Hebb-Williams Maze and the multiple-T maze employing a contingency switch (Blumenthal et al., 2011; Steiner and Redish, 2012; Powell and Redish, 2014). On test sessions, rats rapidly learned the contingency before the switch, and relearned the new contingency after the switch (Figure 1). In order to obtain a measure of flexible behavior on our task, we calculated vicarious-trial-and-error (VTE) at the choice point during test sessions (Muenzinger, 1938; Tolman, 1938; Johnson and Redish, 2007; Papale et al., 2012) by measuring the change of angular velocity of the rat's head (Papale et al., 2012). Before the switch, VTE did not change across laps, contrary to previous results, but after the switch, VTE increased significantly, as the rats were forced to change their behavioral strategy, similar to previous results (Figure 1, Blumenthal et al., 2011; Steiner and Redish, 2012; Powell and Redish, 2014).

Cell Categorization

We recorded a total of 1027 (840 phasic, 187 non-phasic) spike trains from 6 rats over six sessions per rat. Following standard practice, cells were divided into phasic vs. non-phasic cells, reward and non-reward, and maze-responsive and non-maze responsive (Alexander et al., 1986b; Kimura, 1990; Schmitzer-Torbert, 2004). In the aDLS, we

recorded a total of 659 (558 phasic, 101 non-phasic) neurons. In the pDMS, we recorded a total of 368 (265 phasic, 103 non-phasic) neurons. When comparing aDLS to pDMS, proportion of phasic and maze-responsive neurons was not significantly different, but proportion of reward neurons and reward neurons that were also maze-responsive neurons was significantly greater in aDLS (Table 1).

Table 1. Cell Type categorization and percentage of each cell type.

Proportion of Neurons	aDLS Mean ± SEM	pDMS Mean ± SEM
Phasic Neurons	85 ± 4%	72 ± 13%
Reward Neurons	***47 ± 2%**	33 ± 4%
Maze Neurons	47 ± 4%	37 ± 2%
Reward/Maze Neurons	***26 ± 1%**	12 ± 1%

Firing Rate Over Laps

Several previous studies have discovered task-bracketing, in that cell firing increased at the beginning and end of the maze, in the aDLS over sessions (Jog, 1999; Barnes et al., 2005; Thorn et al., 2010; Smith and Graybiel, 2014), but other studies have not found this effect. We investigated whether we would observe a similar effect within session when measured across laps, and if this change would develop before the contingency switch and then re-occur after the switch. To investigate this possibility, we used a similar method as Thorn et al. (2010) and created a linearized representation of each lap by

identifying key points on the maze and measuring firing rate ± one second around each maze event (see methods for details).

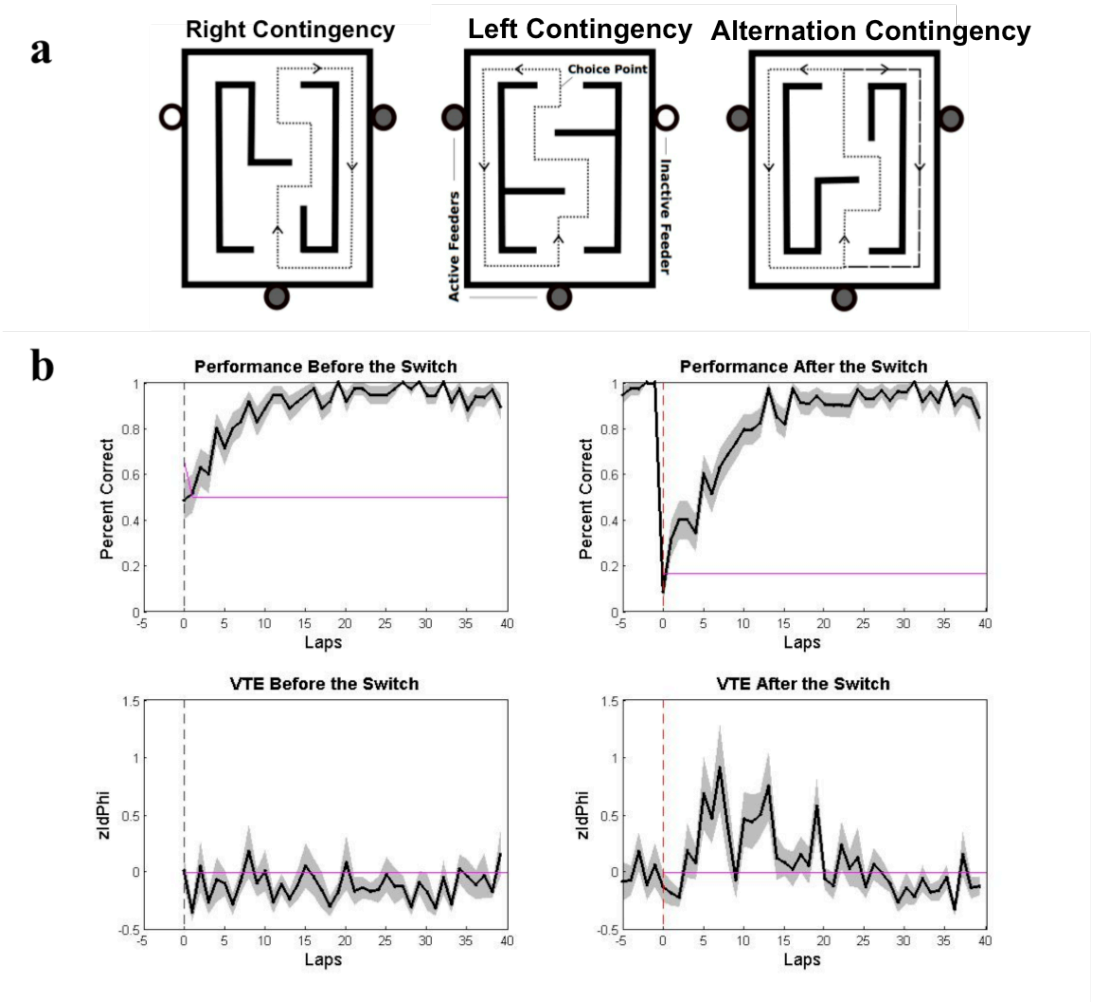


Figure 1. (a) Schematic of the Hebb-Williams Maze with examples of different inner maze components and different contingencies. (b) Performance on maze as measured by the number of errors made each lap before (top left) and after (top right) the switch. Purple lines mark chance (top left) and perseveration rate (top right). Vicarious-trial-and-error (VTE) as measured by zIdPhi (a measure of angular head movement) before (bottom left) and after (bottom right) the switch. Black dotted lines indicate the beginning of the session, and red dotted lines indicate the switch in contingency lap. Purple lines mark zero.

Laps were separated into VTE and non-VTE laps, and on non-VTE laps we observed the overall firing rate in the aDLS increase across laps before the switch, reset then increase again after the switch (Figure 2-side panels). The color plot of firing rate across laps at all maze events indicated that aDLS firing rate increased primarily at the point of the maze that signaled the beginning and end of each lap (Figure 2). In contrast, dynamical firing rates were not as apparent in the pDMS, although pDMS neurons tended to fire around mid-maze events, such as the choice point, feeder click, feeder entry, and return arm.

To investigate specific task-bracketing effects across laps, a measure was used similar to that used by Smith and Graybiel (2014), taking the mean firing rate at the part of the maze signaling the end and beginning of each lap minus the mean firing rate over the rest of the maze (Figure 3). We measured the task-bracket index on Early Laps (1-15) and Late Laps (16-30) before and after the switch. Before the switch, a 2-way ANOVA (Region (aDLS vs pDMS) X Laps (Early Laps vs Late Laps)) was computed, and a significant main effect of Region ($F(1) = 6$, $p = 0.0144$) and a significant interaction of Region X Laps ($F(1) = 4.43$, $p = 0.0354$) was found. Multiple comparisons with Bonferonni-corrected-paired ttests showed that aDLS Late Laps was significantly greater than aDLS Early Laps ($p < 0.025$), showing that task-bracketing in the aDLS increased across laps, and aDLS Late Laps was significantly greater than pDMS Late Laps ($p < 0.025$), providing evidence that aDLS increased and pDMS decreased across laps (Figure 3).

After the switch, a 2-way ANOVA (Region (aDLS vs pDMS) X Laps (Early Laps vs Late Laps) was computed, and a significant main effect of Region ($F(1) = 5.57$, $p = 0.0184$) was found. Although no significant interaction was found, results showed that aDLS Early Laps after the switch was significantly less than aDLS Late Laps before switch ($p < 0.05$, Figure 3) and then increased (though not significantly) again on late laps, nearly reflecting the results before the switch. Multiple comparisons with Bonferonni-corrected paired-ttests showed that aDLS Late Laps was significantly greater than pDMS Late Laps ($p < 0.025$, Figure 3).

The task-bracketing index measure was also used to investigate whether distinct rates of firing occurred and developed at other areas of the maze and whether this was a region-specific effect. Although there were no additional significant interactions, such that the rate of firing did not develop or decline over laps, there were several instances where a significant effect of Region was found, such that overall firing rate was higher in either aDLS or pDMS, and dependent upon maze location (Figure 4).

Specifically, a main effect of Region was found at the navigation sequence, where pDMS firing rate was greater than aDLS, before ($F(1) = 9.71$, $p = 0.0019$) and after ($F(1) = 4.52$, $p = 0.0336$) the switch, at the choice point before ($F(1) = 10.32$, $p = 0.0013$) and after ($F(1) = 30.74$, $p < 0.0001$), and at the feeder entry before ($F(1) = 48.05$, $p < 0.0001$) and after ($F(1) = 77.35$, $p < 0.0001$). In addition to the overall elevated firing of aDLS compared to pDMS at the end zone (see above), a main effect of Region was found at the feeder exit before ($F(1) = 77.49$, $p < 0.0001$) and after ($F(1) = 96.41$, $p < 0.0001$) the switch, where aDLS firing rate was higher than pDMS task bracketing. The importance

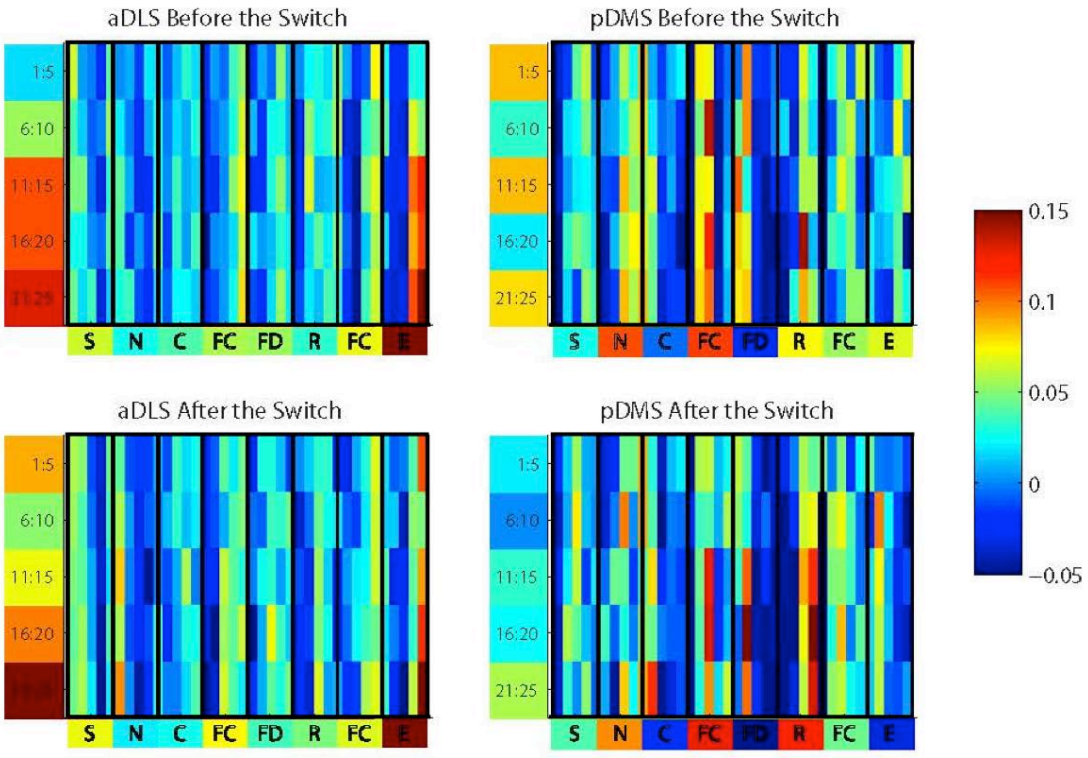


Figure 2. Firing rate over five lap bins centered on specific maze locations in aDLS before (top left) and after (top right) the switch and in pDMS before (bottom left) and after (bottom right) the switch. Side bar of each plot indicates lap bin and shows average firing rate for that bin. Bottom bar for each plot indicates maze location and average firing rate for each location. S = start zone. N = navigation sequence. C = choice point. FC = feeder click. FD = side feeder. R = return arm. E = end zone.

of these additional findings indicates that, while pDMS showed elevated firing at events in the middle of the maze and aDLS showed elevated firing at the feeder exit and end zone, only at the end zone (the maze location that marked the beginning and end of each lap) was there a development of aDLS task bracketing, results similar to findings from previous studies (Thorn et al., 2010; Smith and Graybiel, 2013).

Tuning Curves

Previous research reported dorsal striatal neurons representing parameters of the task necessary in order to obtain reward. For example, on a spatial task, dorsal striatal neurons represented space (Schmitzer-Torbert, 2004; Schmitzer-Torbert and Redish, 2008). In contrast, on a non-spatial task or a cued task, dorsal striatal neurons did not represent space, rather they represented specific actions (Berke et al., 2009) or sequences (Schmitzer-Torbert and Redish, 2008; Thorn et al., 2010) on the task. We were interested in how aDLS and pDMS neurons might differ in their representation of a spatial task, such as the HWM, used in the present study.

In order to better understand how aDLS and pDMS neurons represented spatial information on our task, we created a linearized representation of the maze over laps, and tuning curves were measured. Previous studies have implicated the role of the anterior DMS (aDMS) in strategy switching, such as facilitating and maintaining the acquisition of new behavior strategies (for review, see Ragozzino, 2007), but one study did not find neural correlates of a strategy change (Kimchi and Laubach, 2009). Yin and Knowlton (2004) found that pDMS, but not aDMS, was involved with goal-directed behavior. Thus, we investigated whether *pDMS* would underlie a change in strategy, and the contingency switch in the present experiment provided an opportunity to uncover potential underlying neural correlates of a change in strategy.

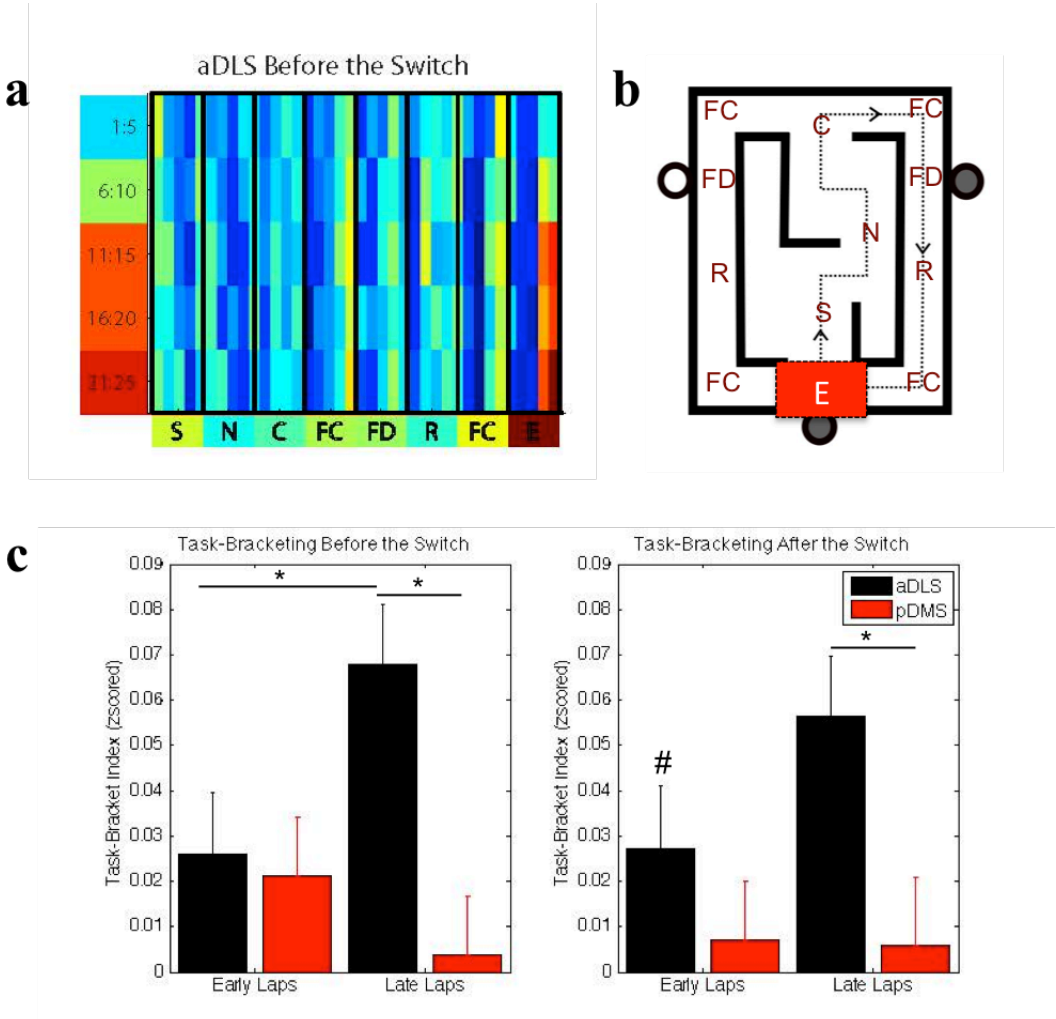


Figure 3. (a) Firing rate plot taken from figure 2 to show an example plot where the task-bracketing measure was taken. (b) Schematic of the maze with maze locations indicated. (c) Task-bracketing index in aDLS (black) and pDMS (red) before (left) and after (right) the switch. S = start zone. N = navigation sequence. C = choice point. FC = feeder click. FD = side feeder. R = return arm. E = end zone. * indicates significant difference. # indicates significant difference of aDLS late laps before the switch to aDLS early laps after the switch.

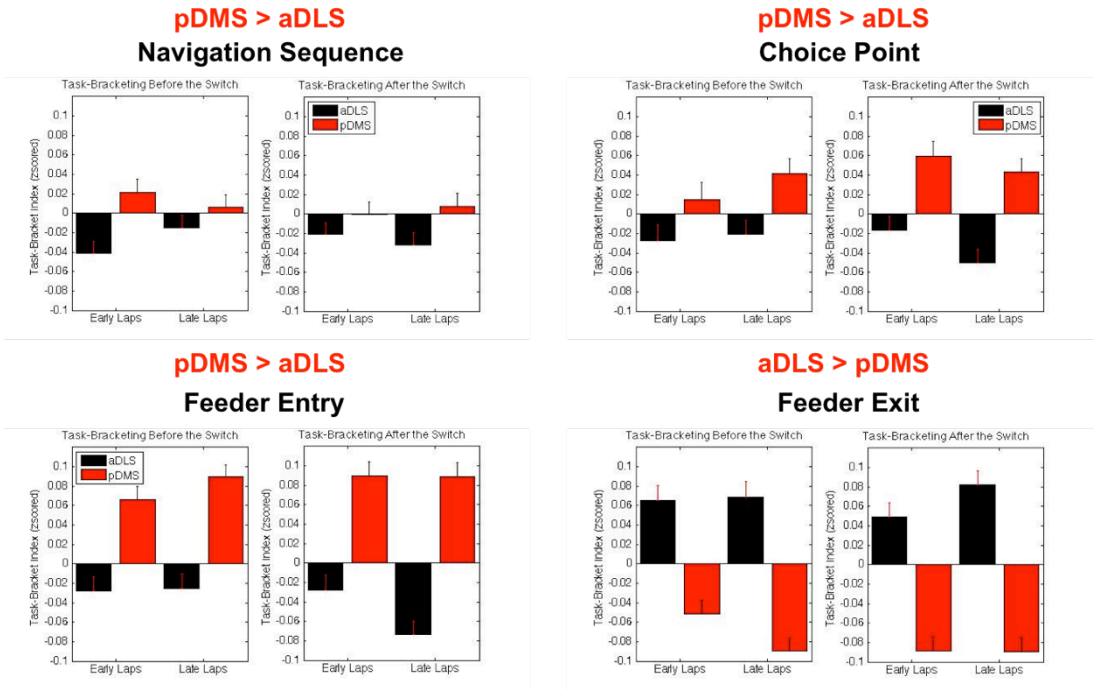


Figure 4. Task-bracketing index at other maze locations. pDMS task-bracketing was higher overall at the navigation sequence (top left), the choice point (top right), and the feeder entry (bottom left). aDLS task bracketing was higher overall at the feeder exit (bottom right).

Close inspection of individual tuning curve examples revealed striking differences in how aDLS and pDMS were representing space on the task. Individual examples indicated that aDLS neurons were representing specific maze locations on specific sides. For example, before the contingency switch, aDLS neurons would be tuned to left laps at the feeder location on a left contingency, and after the switch to an alternation contingency, aDLS neurons would still be tuned to left laps, but not right laps, at the feeder and choice point locations (Figure 5). This indicated that aDLS neurons were representing side of the maze differently, possibly coinciding with specific motor actions, such as taking a left turn or arriving at the left feeder.

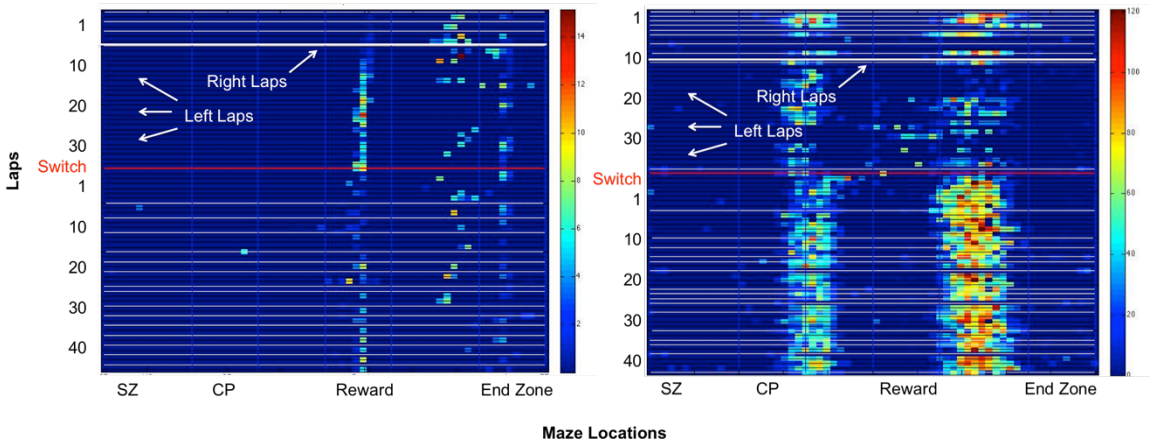


Figure 5. Example of an aDLS (left) neuron that fired almost exclusively on left laps, increasing firing rate across laps before and after the switch. Example of an pDMS (right) neuron that had minimal firing before the switch but increased firing dramatically after the switch. White horizontal lines mark the right laps. Dark horizontal lines mark the left laps. The red horizontal line marks the switch lap. Blue vertical lines mark specific maze locations. SZ = start zone. CP = choice point. Reward = feeder.

In contrast, there were several examples of tuning curves in the pDMS appearing to change as a function of contingency switch. For example, before the switch, pDMS neurons would be moderately tuned to the choice point exit and return arm. After the switch, pDMS tuning curve representation visibly increased and remained elevated until the end of the session (Figure 5). The observed individual examples indicated that pDMS neurons were representing information differently before and after the switch.

Correlation Analyses

To investigate the possibility that aDLS neurons were changing their firing rate on left vs. right laps than pDMS, and pDMS were changing their firing rate more on before vs. after switch laps, we correlated average firing rate on left vs. right laps and on before vs. after the switch laps. Specifically, we took the average of all left laps and an average of all

right laps for each aDLS and each pDMS neuron for each rat and for each test session across the maze to preserve maze locations. We did this for all before and all before laps, as well. We then took a correlation measure for averaged left laps to averaged right laps as well as for averaged before switch laps to averaged after switch laps for every aDLS and every pDMS neuron. To control for lap side and contingency switch, we examined left laps before the switch vs. left laps after the switch by left vs. right laps before and after the switch (Figure 6), and we examined right laps before the switch vs. right laps after the switch by left vs. right laps before and after the switch (Figure 7).

Dorsolateral Striatum Represents Left and Right Laps Differently

If neurons were changing their firing rate dependent upon lap side, we would expect to find less correlation comparing averages of left vs. right laps. Indeed, results indicated that aDLS neurons changed their firing rate on left vs. right laps more so than pDMS neurons. Consistently, total number of correlations for individual neurons closer to zero was greater in aDLS than pDMS. This was observed for aDLS in all conditions (Figures 6 and 7).

For the left vs right laps analysis, a greater number or neurons closer to zero (or lower total correlation) in aDLS than pDMS means that aDLS tuning curves were more sensitive to lap side compared to pDMS. This supports our hypothesis that aDLS neurons were, indeed, more rigid in their representation, often responding at feeders or turns on one or the other side of the maze. This rigid representation may be a reflection of the habit-based nature of the aDLS where neurons tend to respond to particular motor actions

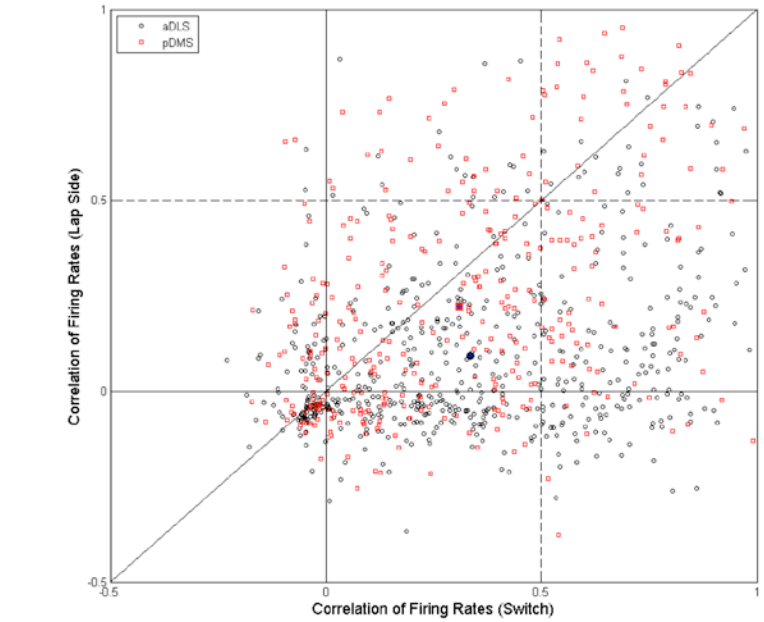
(Carelli and West, 1991; Cho and West, 1997; Miyachi et al., 2002; Yin et al., 2009), such as taking a left turn or arriving at the left feeder.

These results indicate that pDMS cells responded more to similar locations on the maze on both sides than aDLS. That pDMS were tuned to similar locations on the maze regardless of lap side may reflect the more goal-directed nature of the pDMS (Miyachi et al., 2002; Yin and Knowlton, 2004; Yin et al., 2005a; Lex and Hauber, 2010; Lee et al., 2014), with pDMS neurons responding more to sequences, such as taking any turn or arriving at any feeder or any action-association situation that is required to learn in order to obtain reward.

Dorsomedial Striatum Represents Information Differently Before and After the Switch

If neurons altered their firing rate before and after the contingency switch, we would expect to observe a greater number of neurons showing less correlated activity comparing averaged before vs. after switch laps. Indeed, results indicated that more pDMS neurons had less correlated activity than aDLS neurons in all of the conditions (Figures 6 and 7).

To determine whether there were significant differences between aDLS and pDMS neuronal distributions, we conducted a bootstrap analysis and compared differences of actual distances from mean aDLS x y coordinates to pDMS x y coordinates to randomly sampled mean distances in each of the four conditions (see methods). For each of the four conditions, the actual distances between aDLS and pDMS mean x y coordinates were significantly greater than the randomly sampled distribution of distances ($p < 0.001$). To better visualize the differences between the means and the



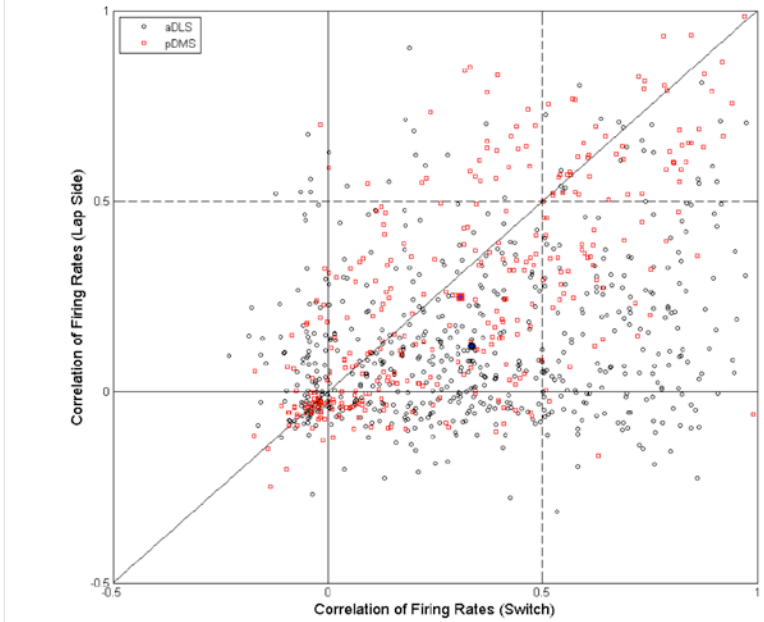


Figure 6. Scatterplots of correlations of firing rate for left laps before the switch vs. left laps after switch laps by left vs. right laps before (top) and after (bottom) the switch. aDLS neurons (black squares) were more sensitive to lap side than pDMS neurons (red circles). aDMS neurons were more sensitive to the contingency switch than aDLS neurons. Filled-in squares and circles indicate the mean of correlations of lap side by the mean of correlations of switch.

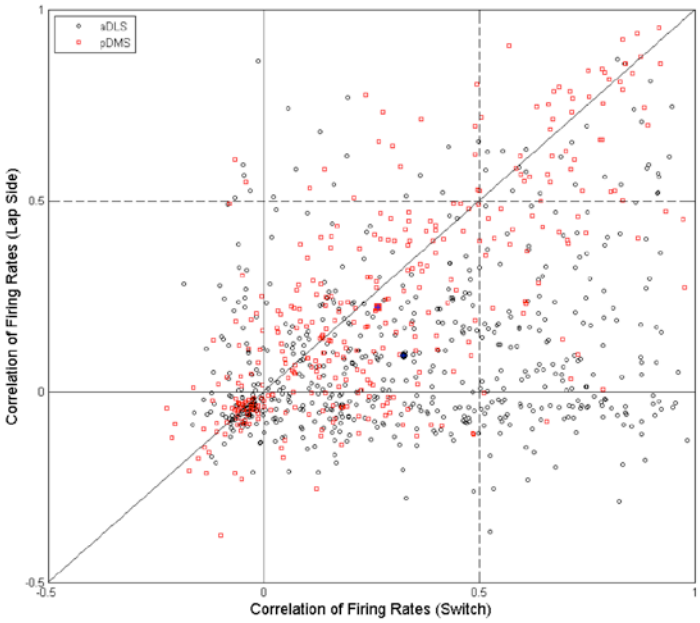
standard errors of the means, all means and standard error bars were plotted on a separate zoomed-in scatterplot (Figure 8).

These results for the pDMS showed that there was less correlation between before and after switch laps in pDMS neurons compared to aDLS neurons. Thus, the results indicate a neural representation of a behavioral strategy change in pDMS neurons, wherein pDMS neurons altered their firing rate dependent upon the switch in contingency, forcing a behavioral strategy change.

Summary of Results

The results from the present study, along with previous studies, indicate that the pDMS contains information related to goal-directed decision making. Representation of spatial information in the pDMS was altered, dependent upon a switch in contingency (Figures 6, 7, and 8), a novel event that forced rats to adjust their behavioral strategy. Neural adaptation of a strategy change occurred even when representing the same spatial actions in the same positions.

Thus, the majority of neurons in pDMS represented a behavioral strategy change. Previous studies that inactivated portions of the dorsal striatum have implicated the DMS as a region important for learning and assessing appropriate actions to optimize reward receipt (for review, see Ragozzino, 2007), for carrying out goal-directed strategies (Yin and Knowlton, 2004) and for action-outcome associations (Yin et al., 2005a; Corbit and Janak, 2010; Shiflett et al., 2010). However, no one had found neural correlates of flexible decision making in the DMS (either aDMS or pDMS) prior to this study.



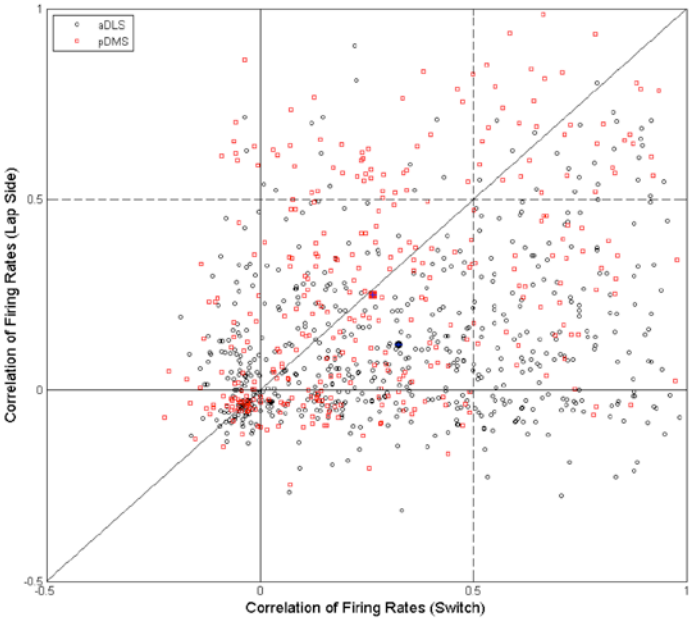


Figure 7. Scatterplots of correlations of firing rate for right laps before the switch vs. right laps after switch laps by left vs. right laps before (top) and after (bottom) the switch. aDLS neurons (black squares) were more sensitive to lap side than pDMS neurons (red circles). aDMS neurons were more sensitive to the contingency switch than aDLS neurons. Filled-in squares and circles indicate the mean of correlations of lap side by the mean of correlations of switch.

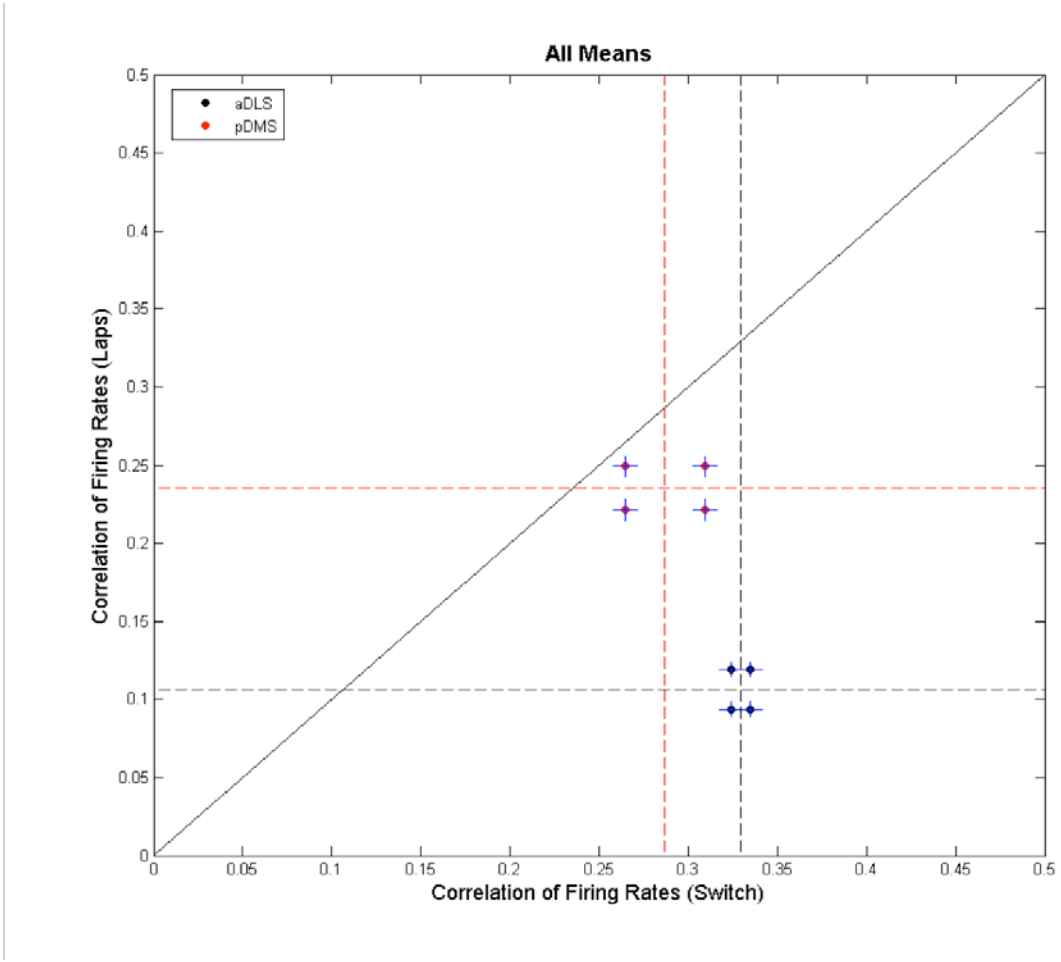


Figure 8. Zoomed in scatterplot of all means from figures 6 and 7 for aDLS (black dots) and pDMS (red dots). Black-dotted lines indicate center position for aDLS means, and red-dotted lines indicate center position for pDMS means. Blue lines on the means indicate the standard error of the means. All means for aDLS neurons were more sensitive to the lap side, while all means for the aDMS neurons were more sensitive to the switch.

The other major finding from this study was the task bracketing-like effects found in the aDLS across laps in the present study, a finding similar to previous studies (Jog, 1999; Barnes et al., 2005; Thorn et al., 2010; Smith and Graybiel, 2013). With a measure of task bracketing (task-bracketing index, compare to Smith and Graybiel (2013)), we found a significant increase of this task-bracketing index at the point on the maze that

signaled the end and beginning of each lap in aDLS across laps before the switch. Task-bracketing index significantly decreased after the switch on early laps and then increased again on late laps (although not significantly). As discussed, previous reports of this phenomenon has been mixed, with several previous studies reporting task bracketing across sessions but not others. This is the first study to observe task bracketing within session, across laps. Task bracketing in pDMS decreased (although not significantly) over laps, such that with aDLS increasing, there was a significant difference between aDLS and pDMS on late laps before and after the switch (Figure 3).

Our results showed that the aDLS contained information related to habit-based behavior, in that the development of aDLS neuronal firing at the beginning and end of each lap coincided with increased experience on the task. In addition, representation of spatial information in the aDLS was more fixed to a specific lap side, a result consistent with a previous study reporting that striatal neurons tended to be involved with egocentricity coding, which means that striatal neurons tended to respond to specific actions, such as going left or right (Berke et al., 2009). Therefore, together, these results suggest that as habit-based behavior increased, aDLS neurons may have been increasingly tied to specific motor actions, such as making a left turn or arriving at the left feeder. This is in contrast to responding to a sequence of actions, such as making any type of turn or arriving at either feeder site, such as appears to be the case with a subset of pDMS reward neurons, in which representations were located at similar parts of the maze around feeder sites.

Recent evidence has shown that pDMS is involved with flexible decision making

(Yin and Knowlton, 2004; Yin et al., 2005a; Yin et al., 2005b; Lex and Hauber, 2010; Shiflett et al., 2010) and contains reward-prediction error neurons that are tied to specific actions (Stalnaker et al., 2012). Results from the present study found that the pDMS contained a majority of non-reward neurons (important for choosing appropriate actions) that were found to be neural correlates of a strategy change. It may be more appropriate to think of the aDLS receiving direction from several brain areas (including the ventral and dorsomedial (especially posterior dorsomedial) striatum).

Some studies implicate that the categorization of the DMS in the actor-critic model may be limiting as to the actual functioning of the DMS, or a new category may be warranted in order to account for the implicated role of the DMS in processing information. For instance, a functional connection between ventral striatum and DLS appears to be important for maintaining rigid, habit-based behaviors, with valuation updated by reward-prediction error information (Belin and Everitt, 2008), and with little input from the DMS. Thus, DMS may function outside the actor-critic framework, or act as a director that dictates the framework the actor and critic function within. It may even act as a filter (a critic of the critic) of information between ventral striatum and DLS until the agent is sufficiently exposed to an environment and habit-based actions predominates behavior.

Previous and present research has shown that aDLS information processing is related to habit-based behavior, and pDMS information processing is related to goal-directed behavior. As an agent learns the environment and the necessary actions in order to obtain reward, corticostriatal control shifts from pre-frontal/orbitofrontal cortical and

pDMS areas to sensory/motor cortical and aDLS areas corresponding with a shift from goal-directed behavior to habit-based behavior. This system of automatizing action sequences is an efficient way for an agent to maximize reward intake while freeing up flexible systems to learn new associations (for review, see Graybiel, 2008). It is highly effective for obtaining natural rewards (e.g. food, water), yet, what happens when drugs of abuse are introduced into this system? What are the alterations that occur and how do these alterations to the system contribute to maladaptive decision-making, such as excessive drug use, typically found in addicted individuals? In the next chapter, I will discuss the implicated role of striatum in addiction and how changes to normal striatal functioning drive maladaptive drug-using behavior.

Chapter 3: Dorsolateral and Dorsomedial Striatum and Addiction

Addiction has been defined as a maladaptive behavior characterized by persistent drug use in the face of negative consequences (American Psychiatric Association, 2000, Diagnostic and Manual of Mental Disorders, 4th ed., text revision). Addiction is considered by many researchers to be a brain disease influencing abnormal behavior (Leshner, 1997; Volkow and Fowler, 2000; Dackis and O'Brien, 2001; Volkow et al., 2003). It is driven by decision-making systems and functions similarly to other behaviors. This is especially true when considering the development of habit-based behaviors. Even though habit-based behavior is difficult to change (discussed in the introduction), individuals still retain at least some agency in their choices, even though choices may seem limited. In the previous chapter, I discussed investigations about normal anterior dorsolateral (DLS), anterior dorsomedial (aDMS), and posterior dorsomedial (pDMS) striatal functioning during decision-making. A foundational understanding of normal brain functioning as related to striatum will be helpful for identifying brain functioning or behavior that deviates from that standard.

Developing habit-based behavior is a normal part of the decision-making process, and habits are instrumental for efficient functioning. They allow us to learn and store repetitive action chains so that we are able to continue learning other action chains (Hikosaka et al., 1995). For instance, think of the many times you have seen a child slowly and purposefully perform some ordinary action that an adult can perform effortlessly, such as tying one's shoes or performing simple arithmetic in one's head. Those small actions may seem big and important to a child but are simple to an adult,

since these actions have been performed all one's life. For the experienced, those actions have been cached and have become habits, requiring little thought, allowing one to focus on other, ostensibly more important matters.

Addiction is similar but also more complex. Addiction can, indeed, be characterized by more automatic, habit-based behaviors (Berridge and Robinson, 1998; Pelloux et al., 2007; Porrino et al., 2007; Belin and Everitt, 2008), but some theories of addiction also imply purposeful, goal-directed behaviors (Koob and Le Moal, 2001; Redish et al., 2008; Bickel et al., 2011). Below I will discuss several ideas for what drives addictive behaviors as well as potential vulnerabilities for increased risk to greater drug-seeking behaviors, with added emphasis on the ideas that involve dorsal striatal functioning.

Shifting Balances

One idea is that with extended drug use there is a change in behavior from more flexible to more automatic decision-making with underlying changes of neural functionality, driving automatic drug-using behavior.

As discussed in the introduction, with any behavior, performing an action repeatedly will make behaviors more automatic. There is a shift from flexible to more habit-based behavior, wherein behavior becomes more rigid, stereotyped, and difficult to devalue (Adams, 1982; Dickinson, 1985). Along with changes of behavior, there is a shift of underlying neural correlates from ventral to dorsal portions of the striatum (Smith-Roe and Kelley, 2000; Ito et al., 2004; O'Doherty, 2004; Porrino et al., 2004; Vanderschuren

et al., 2005; Atallah et al., 2007; Porrino et al., 2007; Graybiel, 2008; Everitt and Robbins, 2013) and corresponding cortical control shifting from pre-frontal and orbitofrontal cortex areas to motor and sensory cortex areas (Izquierdo et al., 2004; Gremel and Costa, 2013; Lucantonio et al., 2014).

However, early ideas about ventral to dorsal striatal control with corresponding cortical inputs were incomplete. New data is emerging that, as discussed in the previous chapter, takes into consideration distinct regions of the dorsal striatum, each with distinct cortical input, driving different types of behavior (Devan et al., 1999; Yin et al., 2004, 2005a, 2006; Thorn et al., 2010; Chapter 2). Thus, instead of neural activity shifting from ventral to dorsal striatum, it might be more accurate to think of this neural shift as more from medial (ventral and dorsomedial) striatum to anterior dorsolateral striatum (aDLS). In this way, associations are eventually cached by aDLS from areas important for learning, such as the posterior dorsomedial striatum (pDMS) and ventral striatum (VS). In fact, a recent study has implicated that pDMS may act as an intermediary between VS and aDLS (Stalnaker et al., 2012).

Therefore, I would posit that striatal control shifts from initial valuation in ventral striatum, ongoing evaluation in the dorsomedial striatum, and eventual storage in dorsolateral striatum. This would be an efficient system of automatizing actions that allows for flexible decision-making governed by the ventral striatum to initially evaluate values associated with actions, then send control to the dorsomedial striatum that would pair these values with actions. After the agent has been exposed to the action-value pair for a sufficient number of trials, pDMS would send control to the aDLS, where cues

regularly associated with specific actions would be recognized quickly, triggering the specific action. Thus, behaviors would be performed with minimal cognitive demand, allowing the ventral striatum and pDMS to make new associations.

If this type of learning is a normal part of functioning, the question is what is unique about addiction. If normal behavior uses this system to efficiently cache actions in order for other associations to be formed, then it is possible that addictive substances somehow exploit this system, perhaps causing actions to be cached more quickly than non-drug rewards (Nelson and Killcross, 2006; Piray et al., 2010). Or, perhaps, drug-seeking actions are cached more slowly than non-drug rewards (Murray et al., 2014), which may allow for action-value associations to be evaluated longer by flexible decision-making systems, increasing the valuation of drugs.

Just as with normal decision making, studies report that drug use is goal-directed at first (Olmstead et al., 2001) and becomes habit-based (Zapata et al., 2010) after enough use (for reviews, see Everitt and Robbins, 2005, 2013), in which the dorsal striatum has been implicated as playing a role in the formation and maintenance of habit-based drug taking. For example, humans and monkeys have significantly altered dopaminergic signaling after extended cocaine use (Porrino et al., 2004; Volkow et al., 2006; Porrino et al., 2007). In addition, inactivation of the aDLS after extensive drug use reduced drug-seeking behavior (Fuchs et al., 2006; Zapata et al., 2010; Corbit et al., 2012; Gremel and Costa, 2013). Together, the evidence of shifting balance from flexible to habit-based decision-making accounts for some of the factors that make up addictive behavior. However, only a subpopulation of users exhibits maladaptive habit-based behavior that is

resistant to punishment. Below, I will discuss the phenomena of compulsivity and the implicated role of striatum.

Compulsivity

In a subset of drug users, behavior becomes compulsive. Compulsive behavior is characterized by continued use of drugs even in the face of negative reinforcement and, as discussed, is one of the hallmarks of addiction (DSM IV TR). Behavior that is compulsive is often difficult to observe, since drugs of abuse have proven to be difficult to devalue by conventional means (e.g., lithium chloride, satiation) (Everitt and Robbins, 2013, Introduction). Using extended access to cocaine (Ahmed and Koob, 1998) and resistance to punishment (Deroche-Gamonet et al., 2004) methods, Robbins and colleagues designed a pre-clinical model of compulsivity. Subjects were given extended access to cocaine, and after extended cocaine use, a subset of subjects exhibited punishment-resistant behaviors (Pelloux et al., 2007). That they found only a subset of subjects that were punishment-resistant was an intriguing finding, since the percentage (~20%) was similar to the percentage of humans who were susceptible to drug abuse (Anthony et al., 1994). A compulsive pre-clinical population provides a potentially useful tool for studying and finding treatments for addiction (Chapters 5 and 6).

I discussed previously that the aDLS is implicated as playing a key role in the development of habit-based behavior (Chapter 2) and habit-based drug-seeking behavior (see above). I also discussed that pDMS is implicated as playing a central role in flexible behavior (Chapter 2) and goal-directed drug-seeking behavior. How do these two areas contribute to compulsive drug taking?

In human subjects, compulsive behavior has been linked to the anterior caudate (analog of rat aDMS), as measured by the occurrence of perseverative errors (persisting with an old strategy when a new one is required to obtain a reward), with activation of the caudate anti-correlated with number of errors made (Ersche et al., 2011). The number of errors and functioning of anterior caudate both improved when subjects were treated with the dopamine 2/3 receptor agonist, pramipexole. As discussed, similar results have been reported in pre-clinical populations, in which inactivating the aDMS impaired reversal learning (Chapter 2). However, in these studies, compared to control rats, aDMS lesioned rats would perform more regressive errors (Ragozzino, 2007), which is correctly changing strategies but then reverting back to the old strategy.

In cortical areas, compulsive behavior and disorders have been linked to orbitofrontal (OFC) and pre-frontal cortices (Rosenberg and Keshavan, 1998; Saxena et al., 1998; Graybiel and Rauch, 2000; Menzies et al., 2008). As noted previously, the DMS receives different input from cortical areas depending on the location on the anterior to posterior axis. For example, pDMS receives input from the OFC, pre-limbic (PrL), and infralimbic area, and the aDMS receives input from the PrL and anterior cingulate cortex (Chapters 1 and 2).

As discussed, inactivation of the OFC impaired reversal learning, causing more perseverative errors (one measure of compulsive behavior). Further, research into rodents that were resistant to punishment (animal model of compulsivity, discussed above) found that the PrL was sufficient and necessary for compulsive cocaine-seeking behaviors. Chen and colleagues (2013) gave rats extended access to cocaine and then exposed them

to punishment. Similar to findings from Pelloux et al. (2007), 20% of the rats were resistant to punishment. They found that the PrL area was hypoactive in these rats, and when they selectively stimulated the PrL, the punishment-resistant rats became punishment sensitive (Chen et al., 2013).

Taken together with the results of reduced caudate correlating with more compulsive behavior, the circuits containing DMS/OFC and DMS/PFC corticostriatal loops appear to be an important functional system for minimizing compulsive behaviors by maintaining goal-directed decision-making.

Adding to this idea is research that reported lesions to the aDLS enhance the effect of punishment (Jonkman et al., 2012). This means that abnormal dorsal striatal functioning in some individuals may create associations that do not allow them to process punishment in the same way as those not prone to compulsive behaviors. In other words, once striatal control shifts from ventral striatum and DMS to aDLS and are sufficiently cached by aDLS, becoming automatized, actions are less affected by punishment. Thus, punishment is more effective while an agent is still engaged in flexible decision-making, likely mediated, in part, by DMS. Once engaged in habit-based decision making, effectiveness of punishment diminishes. Furthermore, data suggested increased functional coupling between the ventral striatum and DLS was correlated with increased drug-seeking behaviors (Belin and Everitt, 2008), making the functional role of DMS even more important.

Noncompensable Dopamine

Shifting neural control from ventral striatum to aDLS likely increases the chance of habit-based drug taking; however, only in a subset of organisms does the behavior become compulsive. This type of behavior, in part, appears to be mediated by a reduced DMS/OFC/PFC corticostriatal circuit and an increased DLS/motor/sensory corticostriatal circuit. How might drugs of abuse facilitate this transition?

One theory as to why drugs of abuse are so addictive is that they are over-valued, because the properties of drugs cause parts of the brain's reward system to overvalue cues associated with the drugs. Thus, any reward capable of producing an overvaluation would potentially be addictive (Redish, 2004). Smokers will often claim that their first cigarette was pleasant but was likely dysphoric (Heishman and Henningfield, 2000). This is because nicotine does not act on the pleasure-producing receptors of the brain (for review, see Berridge and Robinson, 1998) only the motivating receptors, bound to by dopamine (for review, see Wise, 2004)

Schultz and colleagues (1998) discovered, with non-drug cue-reward pairs, dopamine neurons initially fired at the time of reward, but after being paired with a cue for a period of time, the dopamine signal propagated backwards and increased firing at the cue but not at the time of reward receipt. When the reward was not given, dopamine neurons would reduce their firing at the time when the reward was normally to be given. Interestingly, any magnitude change of reward resulted in a corresponding increase or decrease in dopamine firing at the time of reward receipt, signaling that the value of the

reward was more or less than expected, respectively. Appropriately, this has been termed the reward-prediction error signal.

The presence of a reward-prediction error signal has implications for addictive behavior. Take the average smoker, who smokes in multiple environments, each environment with its own set of cues. These cues are paired with the dopamine signal, and after time, multiple cues may be paired with the act of smoking. Once these cue-reward associations are firmly established, the cue will predict a reward by increasing dopamine, but if the smoker does not smoke, there will be a decrease in dopamine firing (a decrease in dopamine has been associated both with dysphoria and disappointment (Koob and Le Moal, 2001)). Difficulties in quitting can partially be accounted for by this phenomena, since the smoker would be barraged by numerous cues that signal the onset of reward and subsequent decrease in dopamine, causing the smoker to intermittently be in a state of withdrawal and crave to smoke. However, this is only part of the story, since static reward values typically cannot be signaled by multiple reward cues, an effect called Kamin blocking (Kamin et al., 1969). Interestingly, drugs of abuse may be different.

Unlike static value non-drug rewards, in which dopmaine is only released when the reward is novel, and dopamine firing shifts from reward to cue, the biochemical properties of drugs of abuse cause dopamine to be released every single time the drug is used (for reviews, see Wise, 2004; Pierce and Kumaresan, 2006). For example, cocaine functions by inhibiting reuptake of dopamine in the synaptic cleft (Ritz et al., 1987). Hypothetically, this continuous release of dopamine each time a drug is used causes a

constant back propagation to stimuli (Redish, 2004). Thus, the value of the stimulus associated with the drug would increase every time the drug is used.

Limited evidence, so far, has suggested that some drugs of abuse do not continue to increase the value of the reward (Marks et al., 2010). A recent report found that sign-trackers (assign more incentive value to the cue that predicts reward rather than the location of the reward) learned from a second cue that the value of a reward was increasing but did not learn from another cue that signaled a decrease in reward value (Bacharach et al., 2014). Sign-trackers have been found to be more susceptible to drug taking (Flagel et al., 2009). In addition, since dopamine is released upon consumption of drugs of abuse, it may also mean that several cue-reward pairs could be made and continue to be made. Thus far, results have been mixed, and Kamin blocking did occur with some drugs of abuse (Panlilio et al., 2007). Other drugs of abuse had the effect of disrupting Kamin blocking when given prior to the compound cue, evidence that the reward signal is driven by changes in dopamine signaling (Crider et al., 1982; O'Tuathaigh et al., 2003). Importantly, one report found that high nicotine-responding rats did not show Kamin blocking (Jaffe et al., 2010).

Thus, noncompensatory dopamine may exist for individuals that are more susceptible to addictive behaviors, highlighting the importance for recognizing individual variability (Chapters 4, 5, and 6). For some, drugs of abuse may become over-valued more than non-drug rewards, and abstinence from drugs may require extinction to take place not just for one cue-reward pair but multiple pairs.

How does the striatum fit into this theory? As discussed in the previous chapter, some theorists proposed that ventral striatum acts as the valuator (or critic) of performed actions, since the reward-error signal has been found in this region, while dorsal striatum performs these actions (as the "actor", see (Daw et al., 2006)). However, a recent study reported discovering a small percentage of DMS fast-firing neurons in rats contained the reward-prediction error signal, but is only observed after the rat takes a specific action (Stalnaker et al., 2012). Thus, both the ventral striatum and DMS may be important for the assessment of actions and update to the current framework of reward valuation.

Interestingly, drugs of abuse would affect both the critic and the actor, so that multiple stimuli in the environment could release a specific action by DLS, such as drug seeking or drug taking. In addition, the consumption of drugs would affect the valuation of the drug, so that the ventral striatum (and perhaps the DMS) continues to associate increased value with the drug with continued use. Finally, new cues would continue to be associated with the drug. As discussed, preliminary evidence suggests that noncompensatory dopamine occurs with a subset of drug user. Perhaps, then, this very nature of drugs of abuse is what makes them so addictive for a subset of drug users.

Incentive Salience

Noncompensible dopamine theory can be tied in with the incentive salience theory, which suggests that dopamine signals assign motivating value to cues paired with rewards (Berridge and Robinson, 1998). In contrast to hedonic theory, which states that addiction is driven by dopamine signaling, which the theory states is associated with a pleasurable subjective experience (Volkow et al., 2003; Wise, 2004; Volkow et al., 2004)

incentive salience, instead, separates “liking” from “wanting”. Wanting is thought to be caused by dopamine signaling after a reward is initially liked (caused by opioid signaling). For instance, dopamine depletion results in rats less motivated to pursuing reward but would still consume the reward (and like it) when physically moved to the reward site (Berridge and Robinson, 1998). Thus, according to incentive salience theory, rewards would be sought out due to dopamine-inducing cues causing wanting rather than liking.

Interestingly, this idea fits in well with the idea of habitually or even compulsively taking drugs even if the drugs no longer produced the same pleasurable effect as when first consumed. Together with the shifting balances, compulsivity, and noncompensatory dopamine theories, cues associated with drugs might become overvalued and motivation for drug taking elicited by those cues might become more automated with extended drug. Value of the drugs, once automated, might have minimal influence on motivation to take drugs.

It should be noted that Pavlovian decision-making systems have been found to be involved in driving behavior associated with increased incentive salience, affecting probability of approach behaviors to drug cues and drug rewards (Toates, 1986; Berridge and Robinson, 1998, for other references, see Introduction, van der Meer et al., 2012, and Redish 2013).

Auxiliary Ideas (less implications for striatal dysfunction)

Impulsivity has been associated with compulsivity, in that more impulsive individuals also tend to display compulsive behaviors (Robbins et al., 2012). However, impulsivity is

thought to be a maladaptive behavior early in drug using (Torregrossa et al., 2008) and is characterized by actions taken with little foresight for future consequences (Madden and Bickel, 2010) The underlying correlates are thought to involve a lack of inhibitory control by PFC. For example, studies have reported that lesions to the PFC increased impulsive behavior in rats, and reduced PFC activity has been correlated with impulsivity in humans (Bickel et al., 2007). Orbitofrontal cortex has been associated with impulsivity, as well, though its role appears to be more complex. Lesions to OFC have caused rats to increase the delay to reward beyond the typical threshold in which control rats would switch to the non-delayed but lower value reward (Winstanley et al., 2004). Given that OFC impairment caused deficits in reversal learning, the increased delays could be a result of the rats' inability to associate their actions with future outcomes. Because DMS receives input from PFC and OFC, the DMS may also have a role in impulsive behaviors.

Finally, according to the opponent processes theory, some individuals may abuse drugs due to changes in homeostatic set points (Koob et al., 1998; Koob and Le Moal, 2001). In this case, stopping use of drugs causes severe withdrawal, especially after long-term use. According to the opponent process theory, over a long enough time of drug use, there is a heightened negative response to drugs and a dampened positive effect (i.e., allostatic state). The theory posits that the positive process begins immediately upon consuming drugs but is offset by the onset of the delayed but longer acting negative process (Solomon and Corbit, 1974). Thus, once the positive effect diminishes, the negative effect still persists, causing withdrawal, and influencing an individual's choice to actively seek out more drugs in an attempt to negate the negative effect (Koob and Le

Moal, 2001). Biologically, this effect is thought to be caused by a dampening of dopaminergic (and opiate) signaling and increased stress-inducing neuropeptides (e.g., corticotropin releasing factor, or CRF). This allostatic state causes a negative emotional state that may be alleviated temporarily by using drugs (Koob and Le Moal, 2001).

Above, I briefly mentioned studies that investigated differences between compulsive and non-compulsive rodents. Discovering individual differences in animal models of drug-seeking behavior has been extremely valuable for understanding how and why certain populations are more vulnerable to drug addiction than others, and they provide means to match effective treatment with specific characteristics. Several animal models have been developed to help understand individual differences of drug addiction, and I will spend the next two chapters discussing individual differences in animal models of addiction (Chapter 4) and matching effective treatments to specific characteristics (Chapter 5).

Chapter 4: Rodent Models of Addiction: Two Theories

The previous chapter discussed addiction broadly, as well as the implicated role of striatum, particularly the dorsal striatum, in different theories of addiction. This chapter will focus on animal models that differ in their drug-seeking behaviors, namely rodent models that self-administer cocaine relatively to a greater or lesser degree. A potential role of striatal functioning in driving differential drug-seeking behavior will also be considered.

Rodent models of drug addiction are useful in the effort to investigate different aspects of drug addiction, especially models that use self-administration. Animals are fitted with jugular catheters and trained to execute an action to receive intravenous drug. This method has permitted researchers to observe and measure several phases of drug use which models human drug use, such as initiation of drug use (acquisition), daily use of drugs (maintenance), increased drug consumption (escalation), forced abstinence (extinction), and relapse (reinstatement).

Utilization of the self-administration method has led to the discovery of individual behavioral traits that are predictors of more or less drug use. Several rodent models of addiction with individual variability of drug use have emerged in the last several years. These include high and low novelty-seeking animals (Piazza et al., 1989; Hooks et al., 1991; Davis et al., 2008; Flagel et al., 2010), Lewis and Fischer 344 rats (Kosten et al., 1994, 1997), high and low cocaine-responding animals (see Yamamoto et al., 2013), high (HiI) and low (LoI) impulsive animals (Perry et al., 2005, 2008; Anker et al., 2009; Economidou et al., 2009; Molander et al., 2011), and high (HiS) and low (LoS) saccharin

intake animals (for review, see Carroll et al., 2008). Each animal model displays high and low drug-seeking behavior; typically the “high” characteristic reflects a propensity for relatively more drug seeking. For example high novelty (Piazza et al., 1989), HiI (Perry et al., 2005), and HiS (Perry et al., 2007), acquire drug self-administration more quickly and in greater numbers. High impulsive (Anker et al., 2009) and HiS (Perry et al., 2006) animals escalate cocaine intake to a greater extent. High impulsive (Perry et al., 2008) and HiS (Perry et al., 2006) rats reinstate more to a lever previously paired with cocaine, compared to their counterparts.

Different tasks have been established that measure impulsivity in animal models, such as the 5-choice serial reaction time (5-CSRT, (Robbins, 2002)) task and the delay-discounting (DD, (Richards et al., 1997)) task. In the 5-CSRT task, animals are trained to wait for a brief light cue (500 ms) above one of five different food ports, and by attending to and choosing the correct food port, animals gain access to a reward. Animals that incorrectly attend to the brief cue or prematurely make a response are thought to be more impulsive. These differential phenotypes have been used to study individual differences in drug-seeking, and, as noted above, studies have found that the more impulsive rats acquire cocaine more quickly and reinstate more to cocaine than less impulsive rats .

The DD task consists of a two-lever choice; one lever delivers a small reward immediately, and the other lever delivers a large reward after an adjusting delay. Choosing the immediate side decreases the delay and choosing the delayed side increases the delay, allowing subjects to adjust the wait time for the large reward to their preferred delay. Studies have shown a bimodal distribution of animals that prefer a lower delay

(HiI), considered more impulsive, and animals that prefer a higher delay (LoI), considered less impulsive. As noted above, the animals considered more impulsive exhibited greater drug-seeking behavior than those considered less impulsive.

I will be referring to both of these animal models of impulsivity. In order to avoid confusion, I simply note when impulsivity has been measured by the 5-CSRT task and use an abbreviation (HiI/LoI) to refer to those considered more or less impulsive on the DD task.

Interbreeding of differential phenotypes has been important for showing that distinct genetic traits exist that predict drug-seeking behavior prior to exposure to drugs of abuse. Researchers have bred rats with specific characteristics and interbred them in an attempt to carry on a genetic variant and create more extreme phenotypes. One of these consists of differing phenotypes that prefer relatively more (HiS) or less (LoS) of a sweet liquid reward (water + saccharin). Over the years, as mentioned, these animals have reliably exhibited differing cocaine self-administration behavior, with HiS animals showing greater cocaine-seeking behaviors than LoS rats.

Because of the similarities of drug-seeking behaviors in different rodent phenotypes with unique behavioral attributes (e.g. high or low degrees of impulsivity or high or low sweet-preference), it is tempting to conceive of similar driving mechanisms that explain behavior. I will explore two of these possibilities.

Theory 1: Animals that display less drug-seeking behaviors are more sensitive to aversive events

More or less drug-seeking behavior could be modulated by a differential subjective experience. One idea is that low drug-seeking animals tend to react more to aversive events than high drug-seeking animals (Carroll et al., 2009). Preliminary evidence partially supports this idea, however reports from studies have been conflicting. Thus, the story appears to be more complex than simple reactivity to aversive events acting as a protective factor against increased drug use. Even so, sensitivity to aversive events may help to predict the degree of compulsiveness. That is, if high drug-seeking phenotypes are less sensitive to punishment than low drug-seeking phenotypes, this might provide important information as to why these phenotypes differ in drug-seeking behaviors.

A number of studies report a general difference in phenotypic reactivity to aversive events, but there are also mixed results, dependent upon the specific animal model and task. For example, compared to HiS (more drug seeking) animals, LoS (less drug seeking) animals exhibited greater acoustic startle after ethanol withdrawal (Dess et al., 2005), greater food-deprived locomotor behavior, more baseline acoustic startle, greater stress-induced analgesia (Dess et al., 2000), and greater reactivity to stress in an open field test (Dess and Minor, 1996). In contrast, LoS rats had less sensitivity to acoustic startle after morphine withdrawal (Radke et al., 2013) and greater novelty-induced locomotor activity (Carroll et al., 2007) compared to HiS rats. Furthermore, HiS male rats showed conditioned place aversion to morphine, while LoS male rats did not (Radke et al., 2013). Similarly, Fischer 344 (less drug seeking) rats exhibited greater HPA axis activity in response to stress and greater emotionality reactivity, while Lewis

(more drug seeking) rats had greater acoustic startle and conditioned taste aversion (for review, see Kosten and Ambrosio, 2002).

Novelty-seeking appears to have the most consistent relationship of drug-use and reactivity to aversive events. Low-novelty (less drug seeking) rats exhibited greater stress reactivity than high novelty-seeking (more drug seeking) rats on several tasks (Dellu et al., 1996; Kabbaj et al., 2000; Stead et al., 2006). Thus, it may be more appropriate to label novelty seekers as novelty-avoiding (low drug seekers) and not novelty-avoiding (high drug seekers) rats.

Impulsivity had only a minimal association with aversive events. For example, researchers found that rats selected as high impulsive or low impulsive (selected by the 5-CSRT task) did not vary in stress reactivity (Molander et al., 2011). HiI and LoI rats did not differ in stress-induced reinstatement responses on a lever previously paired with cocaine (Regier et al., 2014). Additionally, withdrawal from various drugs of abuse and the anxiogenic substance yohimbine actually increased impulsivity (Dalley et al., 2006; Sun et al., 2010).

Finally, male (less drug seeking) subjects exhibited greater withdrawal than female (more drug seeking) rats to morphine (Cicero et al., 2002) and nicotine (Hogle and Curtin, 2006), and high and low-cocaine responders did not differ in levels of corticosterone (Nelson et al., 2010).

To further investigate differential effects of reactivity to aversive events in different phenotypes, we trained females (F) vs Males (M), HiS vs LoS, and HiI vs LoI rats on a cocaine self-administration task for 10 days, then added histamine to the cocaine

solution for 10 days, and then allowed access to cocaine without histamine for another 20 days. Histamine was used, because it is similar to the negative consequences of drug use that may involve emotional and physical distress. It may also mimic neurological pruritus, which sometimes occurs after extended cocaine use and is characterized by aversive, delocalized itching sensations on the skin and throughout the body. In pre-clinical studies with non-human primates, histamine has been effective at reducing drug self-administration (Negus, 2005).

Study: Comparison of Histamine Punishment on Different Phenotypes (First Author of this study was Nathan Holtz. I was a co-author.)

Methods

Subjects

Sixty-three adult rats were used in the present study. Of the total, 20 were female Sprague Dawley rats selectively bred for high (HiS; *n* = 10) or low (LoS; *n* = 10) saccharin intake and 43 were Wistar rats (7 male and 36 female). Twenty-six of the female Wistar rats were screened for high (HiI; *n* = 13) or low (LoI; *n* = 13) impulsivity on a delay discounting procedure. Seven male and 10 female Wistar rats were used to assess sex differences. Prior to experimental sessions, rats were pair-housed in plastic cages and allowed to acclimate for a minimum of 3 days where they had free-access to food (Purina Laboratory Chow, Purina Mills, Minneapolis, MN) and water. All rodent holding rooms were maintained at 24°C and at 40-50% humidity under a constant 12/12-hour light/dark cycle with room lights on at 6:00 am. During the experimental

conditions, female and male rats were food-restricted to 16 g and 20 g of daily food intake, respectively. The experimental protocol #1008A87754 was approved by the University of Minnesota Institutional Care and Use Committee. Experiments were conducted in accordance with the Principles of Laboratory Animal Care, and all facilities were AAALAC accredited. Prior to catheterization, rats were anesthetized with a combination of ketamine (60 mg/kg) and xylazine (10 mg/kg). Atropine (0.15 mL), and doxapram (5 mg/kg) were administered to facilitate respiration under anesthesia. A chronic indwelling polyurethane catheter (MRE-040-S-20, Braintree Scientific, Inc., Braintree, MA) was implanted in the right jugular vein. The other end of the catheter was led subcutaneously to an incision made medial and 1 cm caudal from the scapulae and was connected to the cannula embedded in an infusion harness (Instech Laboratories, Plymouth Meeting, PA). For 3 days after surgery, rats were given heparin (0.2 mL, 50 units/mL, iv) to prevent clotting in the catheter and baytril (2.0 mg/kg, iv) to prevent infection. Catheter patency was assessed every 5-7 days by the administration of a 0.1 mL solution containing ketamine (60 mg/kg), midazolam (3 mg/kg), and saline, and a second catheter was implanted in the left jugular vein if a loss of the righting reflex was not manifest during the catheter patency check.

Drugs

Cocaine HCl was provided by the National Institute on Drug Abuse (Research Triangle Institute, Research Triangle Park, NC), dissolved in a sterile 0.9% saline at a concentration of 1.6 mg cocaine HCl/ 1 mL saline, and refrigerated. The anticoagulant heparin (1 mL heparin/200 mL of saline; 190 USP units of heparin/kg) was added to the

cocaine solution. The cocaine solution was infused at a volume of 0.025 mL/100 g of body weight following a lever press during cocaine self-administration sessions, and the duration of each infusion was 1 sec/100 g of body weight (the average infusion time was 2.5 sec). Histamine (16 mg/mL) was dissolved in sterile 0.9% saline and was added to the cocaine solution during the histamine self-administration phase of the study. This dose was chosen based on unpublished dose-response pilot data from our laboratory.

Apparatus

Custom-made operant conditioning chambers with alternating stainless steel and Plexiglas walls were used to conduct both the delay-discounting and self-administration studies. Each chamber contained slots that allowed for the insertion of stainless steel panels, lights, and operant fixtures. For the delay-discounting procedure, a 45-mg pellet feeder was attached to a pellet-delivery trough (Coulbourn Instruments, Allentown, NJ, USA) on one of the stainless steel panels. There were also two response levers mounted on stainless steel panels 4 cm above the cage floor on either side of the pellet receptacle. Tri-colored (red, green, and yellow) stimulus lights (4.6 W) were located above each lever, and a white house light (4.6 W) was positioned at the top of chamber. Operant conditioning chambers were housed within wooden sound-attenuating boxes and MED-PC software (Med Associates, St. Albans, VT, USA) was used to program experiments and collect data.

Operant conditioning chambers used for the self-administration component were identical to those described above with the exception that the two response levers were located on opposite ends of the chamber and a pellet feeder was not used. Additionally, a

syringe pump (PHM-100; Med Associates Inc., St. Albans, Vermont, USA) containing a 35-mL syringe was located outside of the operant conditioning chamber and was used to deliver response-contingent infusions of cocaine.

Procedure

Selective Breeding for Saccharin Preference

Rats were selectively bred at the University of Minnesota from HiS and LoS lines originating at Occidental College (Los Angeles, California, USA). The HiS and LoS lines were maintained through selectively breeding pairs based on extreme saccharin phenotype scores. Fourteen days following experimental procedures, saccharin phenotype was verified using a saccharin preference test developed by Badia-Elder et al. (1996). In this test, rats had access to a 0.1% saccharin/water solution, and the amount of intake of this solution over 24 h was compared to a 24 h baseline intake of water [saccharin score = ((saccharin mL – water baseline mL) × 100) / body weight]. HiS and LoS saccharin scores were 28.6 (± 2.7 SEM) and 22.6 (± 5.4 SEM), respectively.

Delay-Discounting

A delay discounting procedure was used to screen rats for either high or low impulsivity. Experimental sessions began with the illumination of the house light and lasted for fifteen, 4-trial blocks (maximum of 60 daily trials) or 2 hours, whichever occurred first. During each block, rats had access to 2 operant levers that produced either a small-immediate reinforcer (one 45 mg-food pellet) or a larger delayed reinforcer (three 45 mg-food pellets) following a lever press. Reinforcement availability was signaled by the

illumination of stimulus lights located above the corresponding lever. The first two trials of each block consisted of a forced choice of the small-immediate and large-delayed reinforcers (forced-choice trials) and the last two trials consisted of a free choice between both reinforcers (free-choice trials). The delay for the large reinforcer was initially set at 6 seconds. Responding on the small-immediate and large-delayed levers during subsequent free-choice trials, respectively, decreased or increased the delivery of the delayed reinforcer by one-second increments. The final delay of each session served as the initial delay for the subsequent session. At the end of each session a mean-adjusted delay (MAD) of the free-choice trials was calculated and was used to classify rats as either HiI or LoI. Stability on the delay discounting task was reached once MADs did not differ more than 5 seconds for 5 consecutive days/sessions. Rats with stable MADs $\leq$ 9 sec were classified as HiI rats and those with MADs $\geq$ 13 seconds were classified as LoI rats. Five-day average MADs for HiI and LoI rats were 4.5 (± 0.7 *SEM*) and 33.7 (± 2.4 *SEM*), respectively.

Cocaine Self-Administration

Rats were allowed to recover for 3 days in their designated operant conditioning chambers following catheterization surgery. Subsequently, they were trained to lever press for cocaine infusions (0.4 mg/kg/infusion, i.v.) under an FR 1 schedule of reinforcement during daily 2-hr sessions. All self-administration sessions began at 9:00 am with the illumination of the house light and ended at 11:00 am. During sessions, responses on the left lever (active lever) produced one cocaine infusion and the simultaneous illumination of the stimulus lights located above the lever for the duration

of the infusion. Responses on the other lever (inactive lever) illuminated the stimulus lights above it but had no other programmed consequences. During self-administration training, the active lever was baited with peanut butter (0.5 – 1.0g) and three non-contingent infusions were given at the beginning of each session until stable cocaine self-administration was reached. Once rats reached stability (≥ 25 or more infusions for three sessions and active responses exceeded the amount of inactive responses 2:1) peanut butter and non-contingent infusions were discontinued and rats were allowed to self-administer 0.4 mg/kg cocaine for 10 sessions. This constituted the Pre-Histamine phase. Histamine (4.0 mg/kg/infusion) was then added to the cocaine solution and infusions were monitored for 10 additional sessions, constituting the Histamine phase. Subsequently, the histamine/cocaine solution was removed and rats were allowed to self-administer 0.4 mg/kg cocaine for an additional 20 sessions, constituting the Post-Histamine phase.

Data Analysis

For each subject, infusions were averaged into 5-day blocks (2 pre-histamine blocks, 2 histamine blocks, and 4 post-histamine blocks; see Figures 9, 10 and 11). These blocks of average infusions served as the dependent measure in the present experiments. HiS vs. LoS, HiI vs. LoI, and male vs. female data were collected and analyzed independently as separate experiments. Data were analyzed with a mixed linear model (MIXED procedure) using SAS software (SAS Institute Inc., Cary, NC). Fixed effects in this analysis consisted of phenotype or sex (e.g., HiS vs. LoS; HiI vs. LoI; female vs. male), block (1-8), as well as a phenotype × block or sex × block interactions. Blocks of

infusions were further analyzed using preplanned within- and between-subjects contrasts. The Bonferroni correction was applied to these contrasts and, therefore, considered significant if $p < .0025$. Changes in self-administration were also analyzed during the Histamine and Post-Histamine phases relative to Pre-Histamine cocaine self-administration by computing percent change of average infusions self-administered during blocks 3-8 compared to block 2 ([block 2 infusions - block *x* infusions] / [block 2 infusions] × 100). Percent changes for each Histamine and Post-Histamine block were compared between groups using Student's *t*-tests. Results were considered significant if $p < .05$.

Results of Study 1

HiS vs. LoS

While HiS and LoS rats did not differ in their infusions in the initial cocaine-only phase (Blocks 1 and 2 baseline), the percent reduction from the cocaine-only baseline to when histamine was added was significant, but it did not reveal phenotype differences. However, the recovery of baseline levels when the cocaine-only condition was reinstated differed between HiS and LoS rats during blocks 5 and 6, and specifically, the LoS rats were much slower to recover cocaine self-administration than HiS rats in the post-histamine, cocaine-only phase. Figure 9 shows mean cocaine infusions that were self-administered throughout the experimental procedure. Results from these comparisons indicated main effects for phenotype [$F(1,17) = 20.26$, $p < .001$] and block [$F(7,110) = 26.43$, $p < .0001$], as well as an interaction effect between phenotype and block [$F(7,110) = 3.03$, $p < .01$]. Both HiS and LoS rats self-administered fewer infusions during blocks

3 and 4 (Histamine phase) compared to their respective baseline infusion means ($p <$ .0001). LoS animals also had fewer infusions compared to baseline (block 2BL) during blocks 5 and 6 of the Post-Histamine phase ($p < .0001$). HiS animals self-administered more infusions than LoS animals during blocks 5 ($p < .001$), 6 ($p < .0001$), and 7 ($p <$ .0001). There were no differences between the two groups during the Pre-Histamine phases. When percent reduction in infusions from baseline was compared across groups, there were no significant differences between HiS and LoS rats in the relative reduction of cocaine self-administered during the Histamine phase. However, LoS rats showed a greater percent reduction in infusions than HiS rats during blocks 5 [$t(17) = 2.73$, $p < .05$] and 6 [$t(15) = 4.28$, $p < .001$] of the Post-Histamine phase.

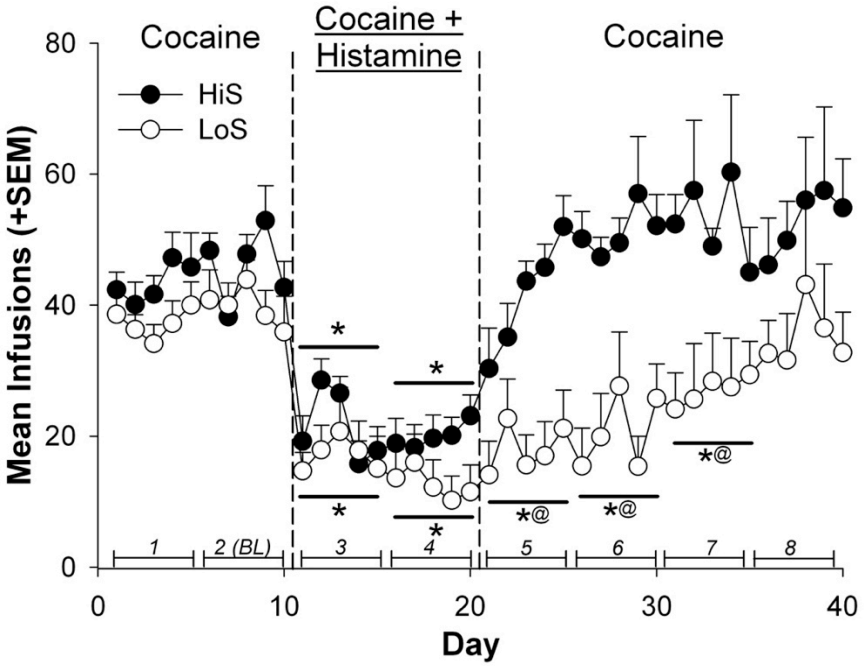


Figure 9: Mean infusions for high saccharin intake rats (HiS, filled circles) and low saccharin intake rats (LoS, open circles) during the pre-punishment (days 1-10), punishment (days 11-20), and post-punishment (days 21-40) phases. Both phenotypes had reduced infusions during punishment, but only LoS rats had reduced infusions during the post-punishment. * indicates significant difference from baseline (BL (days 6-10)). @ indicates significant difference between phenotypes (HiS vs LoS).

HiI vs. LoI

Figure 10 shows cocaine infusions self-administered throughout the experimental phases by HiI and LoI rats. Results from these comparisons indicated a main effect of block [$F(7,152) = 43.75$, $p < .0001$]. Compared to baseline, both HiI and LoI rats self-administered significantly fewer infusions during blocks 3 and 4 of the Histamine phase ($p < .0001$). The HiI and LoI groups did not differ during the Pre-Histamine phase (Blocks 1 and 2 baseline), or in the percent change of infusions between HiI and LoI rats.

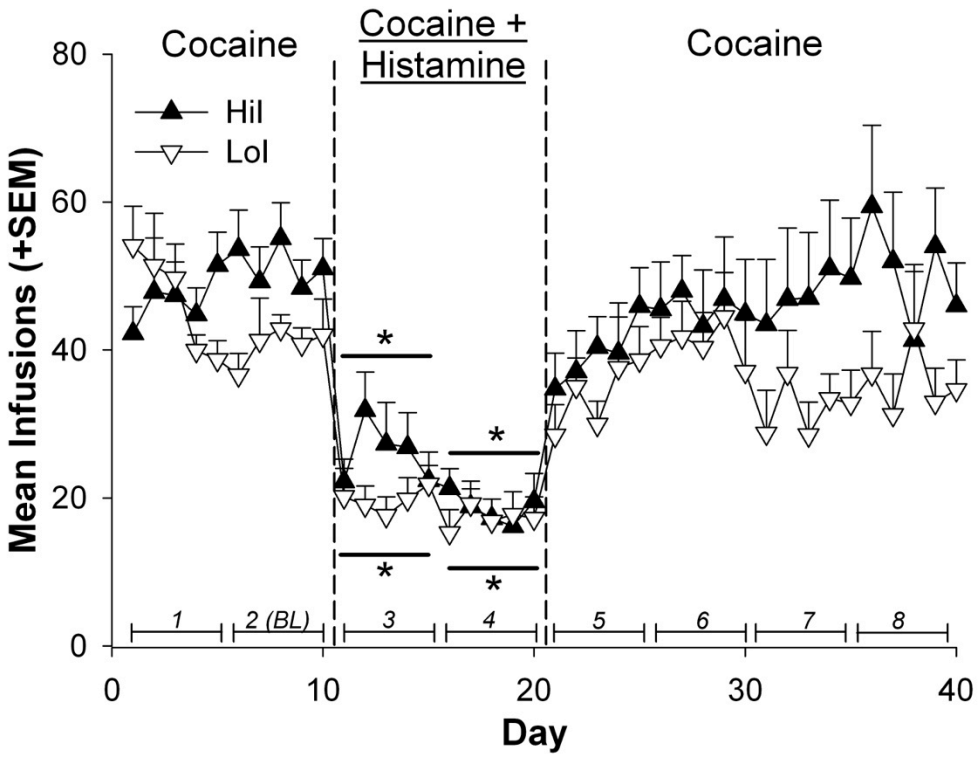


Figure 10: Mean infusions for high impulsive rats (HiI, filled triangles) and low impulsive rats (LoS, open triangles) during the pre-punishment (days 1-10), punishment (days 11-20), and post-punishment (days 21-40) phases. Both phenotypes had reduced infusions during punishment, and both phenotypes quickly increased infusions post-punishment. * indicates significant difference from baseline (BL (days 6-10)).

Females vs. Males

Figure 11 shows cocaine infusions self-administered throughout the experimental procedure by female and male rats. There were no differences between the two groups during the Pre-Histamine phase (Blocks 1 and 2 baseline), nor were there sex differences in percent change of infusions from the pre-histamine to histamine and post-histamine phases. When comparing self-administered infusions, results indicated a main effect of block [$F(7,91) = 12.92$, $p < .0001$]. Compared to baseline (2 BL), female rats self-administered significantly fewer infusions during blocks 3 and 4 of the Histamine phase, as well as block 5 of the Post-Histamine phase ($p < .0001$). Compared to the 2BL, male rats self-administered significantly fewer infusions during block 3 of the Histamine phase ($p < .001$).

Discussion

Results from study 1 showed that punishment with histamine (simulates negative affects of drug use, see introduction) was effective at reducing cocaine self-administration in all phenotypes tested, but only in LoS rats was the effect of punishment prolonged, attenuating a return to baseline in LoS rats after punishment when compared to HiS rats (Figure 9), a result similar to a previous study, wherein animals considered less impulsive (by selection with the 5-CSRT task) showed less reinstatement than high impulsive rats after two phases of punishment cocaine seeking with electric food shock (Economidou et al., 2009). Thus, both high impulsive (5-CSRT task rats) and HiS rats were found to be more punishment resistant than their counterparts Together, the results also may indicate

that HiI and LoI rats in study 1 would have exhibited differential cocaine self-administration after being punished a second time.

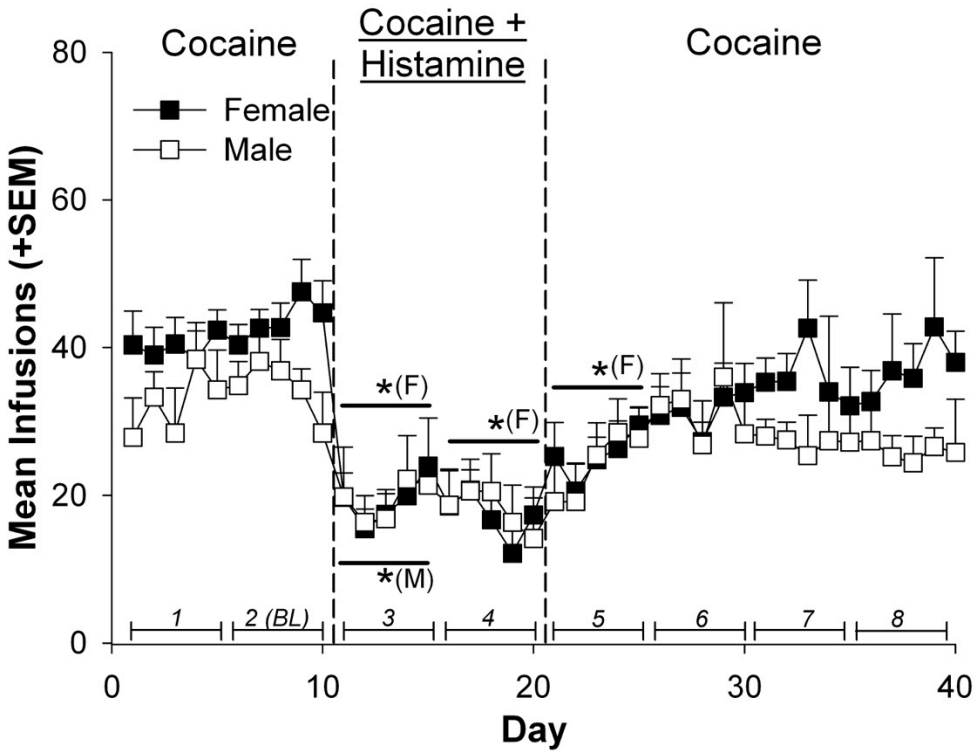


Figure 11: Mean infusions for female rats (F, filled squares) and male rats (M, open squares) during the pre-punishment (days 1-10), punishment (days 11-20), and post-punishment (days 21-40) phases. Both phenotypes had reduced infusions during punishment, and both phenotypes increased infusions post-punishment. * indicates significant difference from baseline (BL (days 6-10)).

Contrary to previous theory (Carroll et al., 2009), the relationship between reactivity to aversive events and proclivity to seek drugs is not straightforward and is dependent upon phenotype of the animal model and type of test used. Results from previous studies and from study 1 indicate that the theory (animals that display less drug-seeking behaviors are more sensitive to aversive events) does not appear to solve the issue of why some animal models take less (or more) drugs compared to others, since it fails to capture the complexities of individual differences in drug seeking. Whereas one

phenotype (novelty seeking) does appear to have an inverse relationship of aversive events and drug seeking, nonetheless most phenotypes (e.g. impulsivity, saccharin intake, Lewis vs. Fischer 344) differ in their response to reactivity of aversive events depending on the type of test administered (see introduction to theory 1).

The results discussed above, along with animal studies that have shown that stress early in life causes increased novelty seeking (Toledo-Rodiriguez and Sandi, 2011), as well as studies in humans and animals that have shown that stress increases drug seeking (Erb et al., 1996; Stewart, 2000; See, 2002; Sinha et al., 2005, 2007; Regier et al., 2014), appear to contradict the broad-sweeping notion that susceptibility to aversive events predicts protection from increased drug-seeking behavior. In fact, even if the theory was true for animals, which does not appear to be the case, it would not translate to human addictive behavior, since individuals who are more susceptible to aversive events (e.g. those with depression and/or anxiety disorders (Eshel and Roiser, 2010)) are 2.7 times more likely to abuse drugs (Regier et al., 1990).

Instead, a more nuanced approach may be more appropriate, in which specific individual characteristics are found and associated with specific behaviors and/or reactivity to specific events. Discovering variations in which individual characteristics predict reactivity to drug taking and other behaviors would have important implications for treatment. For instance, although increased impulsivity does not appear to be positively correlated with stress reactivity but is positively correlated with drug-seeking, withdrawal-induced stress has been shown to increase impulsivity (Dalley et al., 2006). However, the anxiolytic, diazepam, did not reduce impulsivity (Molander et al., 2011).

Thus, the relationship between impulsivity and stress is complex. I will further explore individual responsiveness to treatment in the next chapter, but, first, it is important to consider another theory, which is that similar drug-seeking behavior between divergent drug-seeking phenotypes implicates similar underlying neural correlates of behavior.

Theory 2: Similar drug-seeking behavior is driven by similar underlying neural correlates

Because a number of different animal models display similar drug-taking behavior, an interesting question is whether these animals share similar neural characteristics that drive behavioral outputs. Research into underlying causes of addiction has had a tendency to focus on the dopamine system (for reviews, see Koob and Bloom, 1988; Wise, 1996; Nestler, 2004; Volkow et al., 2004). As discussed in chapters 2 and 3, dopamine is vital for decision-making when procuring non-drug and drug rewards, and variations of dopamine functioning have been shown to play a role in driving abnormal behavior, such as excessive drug seeking. Since I have already discussed the striatal role in addiction generally, I will, instead, focus on high and low drug-seeking animals and review studies that have investigated differences in striatal dopamine functioning (using both indirect and direct methods) between these high and low drug-seeking animals and drug-abusing humans.

Previous research has used the expression of immediate early genes, such as c-fos, as a way to observe the cellular response to illicit drugs (Graybiel et al., 1990), and it has been established that induction of c-fos expression in areas receiving heavy dopaminergic input occurs as a result of dopaminergic innervation (Robertson et al., 1989; Young et al.,

1991; Konradi et al., 1996). Furthermore, dopamine receptor functioning has been shown to mediate induction of c-fos expression. For example, loss of the dopamine 1 receptor (D1R) functioning resulted in no c-fos induction (Drago et al., 1996; Moratalla et al., 1996), and dopamine 2/3 receptor (D2R) agonists alone were not sufficient to induce c-fos expression, while D1R agonists were (Capper-Loup et al., 2002). Loss of D2R functioning, on the other hand, merely resulted in a blunted expression of c-fos (Ruskin and Marshall, 1994; Welter et al., 2007).

Investigations of c-fos induction in animal models of drug abuse have shown differences of c-fos induction between high and low drug-seeking animals in the striatum. For example, research reported greater induction of c-fos expression by forced restraint in the dorsal striatum in high novelty-seeking animals (higher drug seeking) compared to low novelty-seeking animals (Kabbaj and Akil, 2001). Similarly, Lewis rats (higher drug seeking) exhibited greater cocaine-induced (with cocaine doses that previously produced a taste aversion in Lewis rats (Glowa et al., 1994)) c-fos expression in ventral striatum (but not dorsal striatum) compared to Fischer 344 rats (Grabus et al., 2004).

In an effort to begin to understand neurobiological differences between HiS and LoS animals and HiI and LoI animals, we investigated c-fos expression in these animals using a dose of cocaine that elicited more reinstatement of drug-seeking behavior in high drug-seeking rats (Perry et al., 2006, 2008, but see Regier et al., 2014). We found that cocaine-induced c-fos expression was greater in LoI (vs. HiI) and LoS (vs. HiS) rats (Regier et al., 2012). Previous research has shown that HiS and LoS rats did not differ in cocaine-elicited locomotor activity from a single injection of this same dose of cocaine

(Carroll et al., 2007). However, a single injection of cocaine was sufficient to produce differences in neural activity in the striatum (Regier et al., 2012), indicating that neural activity may be a more sensitive measure for predictability of divergent phenotypes. In our study, we found that similar patterns of cocaine-induced expression was found in the dorsal striatum in both impulsive and saccharin rats. Differences within phenotypes but similarities between phenotypes of cocaine-induced c-fos expression in the dorsal striatum may exemplify variations in the shift from ventral to dorsal striatal control of learned stimulus-action associations (discussed below).

Induction of c-fos expression by drugs of abuse or other stimuli shows general cellular activation that can be associated with dopaminergic signaling (see above). However, a more direct way to investigate differences of dopaminergic mechanisms is to study dopamine receptor and dopamine transporter availability. A remarkably consistent type of underlying neural correlate is the differential level of dopamine receptor availability in high vs. low drug-seeking animals. For instance, research has reported that dopamine 2/3 receptor (D2R) availability was predictive of higher levels of impulsivity and greater escalation of cocaine intake in rats, wherein high impulsive rats had lower basal levels of D2R (Dalley et al., 2007). Similar findings have been reported in other phenotypes, showing, generally, that rats more prone to higher drug seeking typically had lower amounts of D2 receptors. For example, Lewis (Flores et al., 1998) and high novelty-seeking (Flagel et al., 2010) rats both were reported to have lower D2 receptors compared to Fischer 344 and low novelty-seeking rats, respectively. In mice, reduced D2R functionality caused reduced sensitivity to reward (Welter et al., 2007). Pre-clinical

discoveries of D2R variability in high and low drug-seeking rodents have translational relevance, as well, since human drug abusers have been shown to have lower D2 receptors than non-drug abusers (Volkow et al., 2009). It has yet to be tested whether HiS/LoS or HiI/LoI animals differ in D2R availability, but research into individual differences of drug seeking might benefit from future studies that investigate this possibility.

In contrast to the consistent reports of lower levels of D2R availability predicting greater drug-seeking behavior, reports of dopamine transporter (DAT) related to drug-seeking behavior have been conflicting. For example, although high novelty rats had more DAT availability than low novelty rats (Dietz et al., 2005), Lewis rats had less DAT availability than Fischer 344 rats (Flores et al., 1998). In normal rhesus monkeys (not high or low drug seeking), DAT binding site density was found to decrease initially in response to initial exposure of cocaine, but it increased after chronic and long-term exposure to cocaine in the ventral and dorsal portions of the caudal putamen and caudate (Letchworth et al., 2001).

Availability of D2Rs during progression of increased cocaine intake has been shown to be an important mediating factor of both impulsivity and drug use. For example, studies in monkeys have shown that prolonged cocaine use results in a decrease in D2Rs in most of the striatum, but to a greater extent in dorsal striatum (caudate and putamen in monkeys), relative to control animals (Moore et al., 1998; Nader et al., 2002). Low impulsive animals also showed a typical progressive decrease in D2R levels in both ventral striatum and dorsal striatum, while high impulsive animals showed some decrease

but only after chronic exposure and only significantly in the DLS (Besson et al., 2013), implicating a delayed decrease in availability of D2Rs in the high impulsive rats (Everitt, 2014).

As discussed in previous chapters, corticostriatal control of learned actions shifts from ventral to dorsolateral striatum, and recent data suggests that at least some neurons in the DMS may be involved as an intermediary between ventral and dorsolateral striatum (Stalnaker et al., 2012). Thus, in the striatum, initial evaluation might occur in the ventral striatum, with ongoing evaluation in the DMS, and eventually actions being cached by the DLS. Therefore, the shift may occur from ventral to DMS to DLS, with an added emphasis of DMS needed to maintain flexible decision-making.

A computational model proposed that lower D2R availability (a result consistently found in high drug-seeking animal models) facilitates a transition from ventral striatum to dorsal striatum control of drug-seeking behaviors (Piray et al., 2010), causing habit-based behaviors to manifest themselves more quickly. A recent study tested this transition from ventral to dorsal striatal control in high and low impulsive rats (selected with the 5-CSRT task), the former previously shown to have lower D2R availability. Therefore, if their model was correct, high impulsive rats would exhibit a quicker transition to DLS control. Surprisingly, they found that there was a *delayed* transition from ventral to DLS in high impulsive rats, as demonstrated by dopamine antagonists in the DLS inhibiting drug-seeking responses only in low impulsive animals during the transition phase (after acquisition but before overtraining), a result opposite of what the model predicted. After extensive training on the task, dopamine antagonists in

the DLS inhibited both high and low impulsive animals' drug-seeking responses (Murray et al., 2014).

Interestingly, previous research with cocaine-induced c-fos expression may add nuance to the discussion of shifting control of drug-seeking behavior from ventral striatum to DLS. In the striatum, high acute expression of c-fos induced by cocaine (found in LoS and LoI animals, see above) was found to be predictive of high chronic expression of the dopaminergic inhibitor, dynorphin (Steiner and Gerfen, 1993). Higher chronic dynorphin expression in the dorsal striatum after chronic administration of cocaine might affect the shift from ventral striatum and DMS to DLS, hence from flexible to habit-based decision-making, differently in high and low drug-seeking rats.

One interpretation from the c-fos and D2R results is that a delayed devolution to dorsolateral striatal control might allow for increased valuation and revaluation of drugs and drug cues more so than a quick transition to DLS. This may help to explain previous results with HiI vs LoI and HiS vs LoS rats, wherein the low drug seekers (LoI and LoS rats) would appear to reach a ceiling of drug-seeking responses in acquisition and escalation of cocaine self-administration, while high drug seekers (HiI and HiS rats) would continue to increase drug-seeking responses (Perry et al., 2005, 2007, 2008; Carroll et al., 2008; Anker et al., 2009).

It is possible that delayed transition to DLS control would allow for increased valuation of drugs of abuse (Chapter 3, reward-prediction error signal). In addition, it is thought that, in humans, inhibition of dopamine-stimulus associations in striatum reduce drug seeking (Volkow et al., 2006). Together these ideas could mean that greater

inhibition of dopaminergic signaling by dynorphin in the dorsal striatum striatum might actually have the affect of facilitating the transition from ventral and dorsal striatal control in low drug seekers, while at the same time inhibiting dopamine-stimulus associations. Thus, once cached by DLS in low drug-seeking animals, these associations might be weaker by comparison. In contrast, less inhibitory control in the dorsal striatum and a delayed transition to DLS might allow for high drug seekers to increase drug-seeking behaviors. If these ideas were correct, they may help to explain divergent behavior within unique phenotypes and would have implications for dorsal striatum as a biomarker for treating drug addiction both pharmacologically and behaviorally (Chapter 5).

Underlying neural correlates better explain divergent drug-seeking behavior than reactivity to aversive events.

Although there is no definitive answer for understanding why different phenotypes display similar drug-seeking behaviors, investigations into the differing neurobiology of phenotypes that self-administer drug differently appear to provide a more clear answer as to what drives higher drug-seeking behavior. Part of the problem with trying to discover similar characteristics that drive behavioral outputs in different animal models of drug addiction is that human addiction is a complex problem, and there are a myriad of issues to consider when trying to understand what drives addictive behaviors, and it is unrealistic to define addiction as a singular problem that can be fixed by a singular treatment.

In the next chapter, I will discuss differences in treatment receptivity with a focus on the animal models of addiction discussed in this chapter, especially HiS/LoS and HiI/LoI rats. I will consider the issue of how to address addiction treatment from multiple perspectives, matching efficient evidence-based treatments with specific characteristics, and then discuss the translation of all of these findings to the treatment of human addiction in chapter 6.

Chapter 5: Individual differences in treatment for addiction

In the previous chapters, I discussed how the striatum is involved with decision-making, addiction, and individual differences of drug-seeking behaviors. The discussion thus far has focused on underlying correlates of abnormal drug-seeking behavior. These correlates have implications for appropriate ways to attenuate drug-seeking behavior in animals, and, notably, for treating addiction in humans. In this chapter, I will discuss how different animal models of addiction (discussed in chapter 4, high and low drug-seeking animals) respond to different treatment efforts and then present a new study on the differential effects of treatment on high (HiI) vs. low (LoI) impulsive rats. Previous and present studies have implications for how striatum may play a role in differential response to treatment and for application into human drug-addiction treatment (Chapter 6).

Investigations into the treatment of individual differences have been limited. Studies thus far indicate response to treatment has been more favorable for animals that have exhibited lower drug-seeking behaviors. For example, drug-seeking responses in low saccharin-intake rats (LoS) were attenuated by systemic baclofen (a $GABA_B$ agonist) more so than high saccharin-intake rats (HiS) during long access, whereas, during reinstatement, responses by both HiS and LoS rats were attenuated by baclofen (Holtz and Carroll, 2011). Similarly, responses by LoS, but not HiS rats during long access to cocaine were attenuated by the neurosteroid, progesterone (Anker et al., 2012).

In high impulsive rats (5-CSRT task), atomoxetine has been shown to decrease impulsivity (Ansquer et al., 2014) and drug-seeking behaviors (Economidou et al., 2011), potentially through similar mechanisms (Everitt, 2014).

Beyond this, no one has investigated whether high and low impulsive animals respond differently to treatment. Discovering individual differences in treatment may be extremely important moving forward in the treatment of human individuals. Humans present a far more complex problem, due to different environmental backgrounds, genetic susceptibilities, and biological irregularities, among other factors. Beginning to find individual treatments for animals may be the first step in trying to understand how we can take each individual person on a case-by-case basis in order to maximize treatment efficacy.

We aimed to model human relapse in HiI and LoI animals by training them to stably self-administer cocaine and then expose them to a number of different types of stimuli, such as yohimbine (an anxiogenic drug, Charney et al., 1989; Feltenstein and See, 2006), caffeine, and cocaine after extinction of cocaine self-administration. We then attempted to attenuate drug-seeking responses in reinstatement with allopregnanolone (ALLO, a anxiolytic neurosteroid, Patchev et al., 1996), shown to be effective in reducing drug-seeking previously in female rodents (Anker and Carroll, 2010). We selected HiI and LoI rats on a delay-discounting task for food and then trained the rats on a self-administration paradigm. After acquiring cocaine self-administration, rats maintained steady responding for 12 days, which was extinguished for 16 days. After drug-seeking responses were sufficiently reduced in extinction, drug seeking on a lever previously

paired with cocaine was induced with three different pharmacological manipulations: cocaine (15 mg/kg), yohimbine (2.5 mg/kg), and caffeine (40 mg/kg). We then used allopregnanolone (15 mg/kg) in an effort to reduce cocaine-seeking behavior, and to observe if HiI and LoI rats would differ in their response to treatment.

Study: Differential treatment effect in HiI vs LoI rats

Methods

Subjects

Adult female Wistar rats (total n = 67) were used for this experiment (Table 4). Estrous cycle was not monitored to prevent disruption of cocaine-maintained behavior by repeated vaginal lavage (Walker et al., 2002). Therefore, results can be generalized across all phases of the estrous cycle. After a minimum of 3 days of acclimation following arrival to the laboratory, rats were housed individually in plastic holding cages and moved daily into experimental chambers for delay-discounting testing. During the self-administration period following completion of delay discounting, rats were housed in experimental chambers. For all conditions, rats were housed in temperature (24^{o} C) - and humidity-controlled rooms where there was a 12-hr light/dark cycle (lights on at 6:00 am). Use of these animals for this protocol was approved by the University of Minnesota Institutional Animal Care and Use Committee and was accredited by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC). Recommended principles of animal care were followed (National Research Council).

Delay Discounting

Adult female rats were tested on a delay-discounting task for food in experimental chambers. These chambers were housed within a wooden sound-attenuating enclosure equipped with a ventilation fan. Each had a port for a water bottle for ad libitum water access and a 45-mg pellet feeder (Coulbourn Instrument, Lehigh Valley, PA) that was mounted on a stainless steel wall and attached to a pellet delivery trough. On either side of the pellet feeder were two standard response levers mounted about 4 cm above the cage floor. One lever delivered 1 pellet immediately after a lever press, and the other lever delivered 3 pellets after an adjusting delay. The lever that delivered 1 pellet immediately and the lever that delivered 3 pellets after an adjusting delay alternated each day. The delay was adjusted based on the rat's choices. A lever press on the immediate side decreased the delay by 1 s, and a lever press on the delayed side increased the delay by 1 s.

The rats were tested at the same time each day, 7 days a week, for 2-h sessions or 60 trials (4-trial blocks), whichever came first. Sessions began with a white house light (4.6 W) and one of the sets of tricolored stimulus (situated above both levers) lights being lit. When lit, the stimulus lights indicated that the lever was active. The first two trials consisted of a forced choice on each lever with stimulus lights lit above the correct lever. During the third and fourth trials, stimulus lights above both levers were lit, and trials consisted of a free choice on either lever. At the end of the session, the final delay was recorded and used for the starting delay on the following day. A mean adjusted delay (MAD) score was calculated by taking the total delay divided by the number of free choice trials. Once a rat completed at least 50 trials per session for five days, and the

difference in MAD scores across those five days was no greater than 5 s, an average MAD score across the five days was calculated. Rats were selected as HiI when the average MAD score was < 9 s. Rats were selected as LoI when the average MAD score was > 13 s. Even though it was rare, rats that fell in between 9 and 13 s were excluded from the study. The rationale for these selection criteria was determined by a previous study in which Perry et al. (2005) used several rats and found a bimodal distribution that matched well to this range of numbers. Data collection and experimental programming were controlled by PC computers and MED-PC software (Med Associates, St. Albans, VT).

Cocaine self-administration

After rats were selected as HiI or LoI, they underwent catheterization surgery. They were anesthetized with ketamine (60 mg/kg) and xylazine (10 mg/kg) and received atropine (0.15 mL) and doxapram (5 mg/kg) to facilitate respiration. A chronic indwelling polyurethane catheter (MRE-040-S-20, Braintree Scientific, Inc., Braintree, MA) was implanted in the right jugular vein. The other end of the catheter was led subcutaneously to an incision made medial and 1 cm caudal from the scapulae and was connected to the cannula embedded in an infusion harness (Instech Laboratories, Plymouth Meeting, PA). For 3 days after surgery, rats were given heparin (0.2 mL, 50 units/mL, iv) to prevent clotting in the catheter and baytril (2.0 mg/kg, iv) to prevent infection.

After the 3-day recovery period, rats were trained to self-administer cocaine (0.4 mg/kg) in experimental chambers identical to the delay-discounting procedure, except that instead of a pellet dispenser, there was a holder for a jar containing ground food. In

addition, there was a syringe pump that contained a 30-ml syringe that delivered COC or SAL into the operant chamber via a tether (C31CS; Plastics One, Roanoke, VA), connected on one end to the rat harness and to a swivel (050-0022, Alice King Chatham, Hawthorne, CA) on the other end. The rats were trained under an FR-1 reinforcement schedule during daily 2-h training sessions. A house light automatically turned on at 9 am every morning signaling the beginning of the session. During this time, a lever press on one lever (active lever) resulted in activation of tricolored stimulus lights above the lever and activation of the pump, which delivered intravenous cocaine at a volume of 0.025 mL/100 g body weight (duration = 1 s/100 g). Responses on the other lever (inactive lever) resulted in illumination of stimulus lights above the lever but had no other programmed consequences. Upon onset of self-administration, catheters and tethers were checked daily for leaks with a heparin/saline solution, and every 5-7 days catheters were checked for patency by flushing with an iv solution containing 30 mg/ml ketamine, 1.5 mg/ml midazolam, and saline (0.1 – 0.2 ml, iv). Patency was inferred by loss of righting reflex. If this test failed, rats underwent a second surgery, in which they were implanted with a catheter into the left jugular.

During acquisition, levers were baited with peanut butter and non-contingent infusions were provided to facilitate exploration of the lever eventual responses by rats. Rats were required to earn 25 or more infusions a session and to maintain a 2:1 ratio of active vs. inactive lever responses for 3 consecutive sessions. Once rats could reliably meet these criteria without the lever being baited or non-contingent infusions being provided, they entered the maintenance phase, where their responses and infusions were

monitored and recorded for 12 days. After 12 days of rats maintaining at least 25 infusions per session with a 2:1 active to inactive lever responses, the extinction period began.

The extinction phase lasted for 16 days, and the consequences of lever responding remained identical to maintenance, except that COC was replaced with SAL. After extinction, and for 3 days prior to reinstatement, the stimulus lights, syringe pump, and house light were disconnected. Responses on both levers during this pre-reinstatement period were recorded but had no programmed consequences.

Subsequently, the rats entered into the reinstatement period, where they were divided into 3 separate groups. Lights and pump remained off during this time, and rats received alternating ip injections of saline (SAL) and drug solution (COC, CAFF, or YOH) with pretreatment 30 min before each daily session with either peanut oil (VEH) or ALLO (15 mg/kg). Each drug and VEH or ALLO combination was administered only once, with SAL and VEH combination being administered on days in between the administration of drug and VEH or ALLO combination (Table 1).

Drugs

Cocaine HCL was provided by the National Institute of Drug Abuse (Research Triangle Institute, Research Triangle Park, NC). It was dissolved in sterile 0.9% saline, and the anticoagulant heparin (1 ml heparin/200 ml of saline) was added to prevent thrombin accumulation. Caffeine and ALLO were obtained from Sigma Aldrich (St. Louis, MO) and dissolved in saline (40 mg/ml) and peanut oil (15 mg/kg), respectively. Yohimbine

(2.5 mg/ml) was obtained from Lloyd Laboratories (Shenandoah, IA) and came in an injectable form.

Table 2. Reinstatement groups and order of priming events

Priming Condition	N		Reinstatement Priming Sequence			
Cocaine (COC)	24	*Pretreatment*	VEH	VEH	VEH	ALLO
		Prime	SAL	COC	SAL	COC
Caffeine (CAFF)	24	*Pretreatment*	VEH	VEH	VEH	ALLO
		Prime	SAL	CAFF	SAL	CAFF
Yohimbine (YOH)	19	*Pretreatment*	VEH	VEH	VEH	ALLO
		Prime	SAL	YOH	SAL	YOH

Data Analysis

Primary dependent measures included MADs during the delay discounting task, responses and infusions during maintenance and extinction of self-administration, and responses during reinstatement of cocaine-seeking behavior. Repeated measures were days during maintenance and extinction and injection type during reinstatement. Outlying values within each group that were two standard deviations outside of the mean were excluded from analysis. MADs were compared between HiI and LoI rats using an unpaired 2-tailed Student's t-test. For maintenance and extinction, responses and infusions were averaged across 4-day blocks to reduce variability and the number of post hoc contrasts and were analyzed using a 2-factor mixed ANOVA (phenotype X days). For reinstatement, groups that received different priming injections (e.g., COC, CAFF, YOH) were analyzed separately using 2-factor mixed ANOVA (phenotype X priming

injection, e.g. VEH/COC, ALLO/COC). Dunn's (Bonferroni) procedure was used for post hoc analyses. All analyses were completed using GBStat (Dynamic Microsystems, Silver Spring, MD).

Results

Mean Adjusted Delay

Confirming behavioral phenotype, HiI mean MAD scores (4.71 ± 0.33) were significantly lower than LoI mean MAD scores (20.67 ± 1.44) ($t(65) = 11.17$, $p < 0.0001$).

Maintenance

Mean responses and infusions (4-day blocks) for the cocaine self-administration maintenance period are shown in Table 5. Results of the 2-factor ANOVA revealed no significant main effect of phenotype or day and no significant phenotype X day interaction for responses or infusions.

Extinction

Table 5 displays the mean number of responses (4-day blocks) over the 16-day extinction period. While there was no significant main effect of phenotype or a phenotype X day interaction, there was a significant main effect of day ($F(3,267) = 171.80$, $p < 0.0001$), indicating a decrease in responding over the extinction period for all rats.

Saline infusions (Table 5) over the extinction period were analyzed similarly. Results indicated no significant main effect of phenotype but a significant main effect of day ($F(3,267) = 128.37$, $p < 0.0001$) and phenotype X day interaction ($F(3,267) = 2.74$, p

< 0.05). Post hoc analyses revealed a notable decline in responding in all rats over the extinction period but no differences between the HiI and LoI rats.

Reinstatement

Cocaine-Primed Reinstatement

Cocaine-seeking responses primed by cocaine are displayed in Figure 12. Results indicated no significant main effect of phenotype or phenotype X priming injection interaction, but there was a main effect of priming injection ($F(1,45) = 10.41$, $p < 0.01$). Post-hoc analyses revealed that ALLO significantly reduced COC-primed responding in LoI but not HiI rats.

Table 3. Mean (± SEM) responses and infusions during maintenance and extinction

Phase	Blocks of 4 days	HiI Responses	LoI Responses	HiI Infusions	LoI Infusions
Maintenance	Days 1-4	50.60 ± 1.60	50.07 ± 1.32	40.41 ± 1.03	38.40 ± 0.91
	Days 5-8	49.63 ± 1.91	53.55 ± 1.87	39.79 ± 1.06	41.24 ± 1.14
	Days 9-12	46.52 ± 1.35	53.30 ± 1.76	38.44 ± 1.07	42.79 ± 1.37
Extinction	Days 1-4	29.52 ± 2.27	32.42 ± 2.62	22.38 ± 1.66	26.22 ± 2.06
	Days 5-8	11.28 ± 0.99	8.68 ± 0.80	8.54 ± 1.66	6.38 ± 0.55
	Days 9-12	6.63 ± 0.63	5.4 ± 0.69	4.69 ± 0.70	3.88 ± 0.44
	Days 13-16	3.90 ± 0.43	3.76 ± 0.50	2.75 ± 0.32	2.81 ± 0.36

Caffeine-Primed Reinstatement

Results for CAFF-induced cocaine seeking are displayed in Figure 13. While there was no significant main effect of phenotype and no phenotype X priming injection

interaction, there was a main effect of priming injection ($F(1,43) = 12.49$, $p < 0.01$). Post-hoc analyses revealed that ALLO significantly reduced CAFF-induced responding in LoI but not HiI rats.

Yohimbine-Primed Reinstatement

For yohimbine-induced cocaine-seeking responses, there were no significant main effects of phenotype or priming injection and no phenotype X priming injection interaction (data not shown).

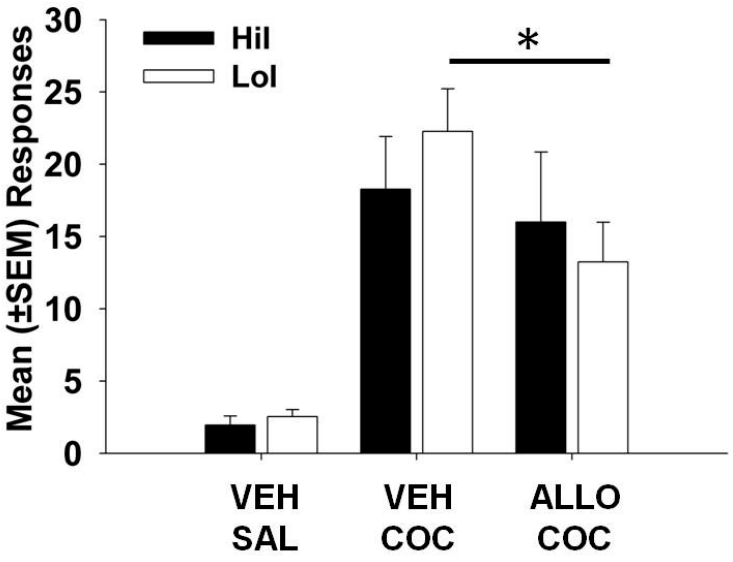


Figure 12. Mean (±SEM) responses for cocaine- (COC) primed reinstatement of cocaine-seeking behavior. Pretreatment with allopregnanolone (ALLO) attenuated cocaine-seeking responses in low (LoI) but not high (HiI) impulsive rats. * indicates significant difference.

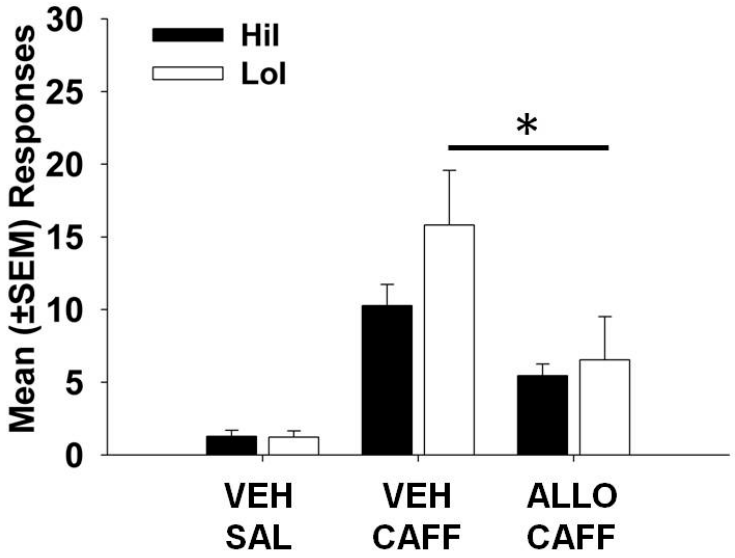


Figure 13. Mean (±SEM) responses for caffeine- (CAFF) primed reinstatement of cocaine-seeking behavior. Pretreatment with allopregnanolone (ALLO) attenuated cocaine-seeking responses in low (LoI) but not high (HiI) impulsive rats. * indicates a significant difference.

Summary

Results from our study indicated that allopregnanolone (ALLO) was more successful in attenuating cocaine- and caffeine-induced reinstatement in low impulsive (LoI) compared to high impulsive (HiI) animals. These results are similar to previous reports of low saccharin-intake (LoS) rats responding more favorably to baclofen (Holtz and Carroll, 2011) and progesterone (Anker et al., 2012) than high saccharin-intake (HiS) rats.

That HiI and LoI rats did not differ in response to yohimbine (an anxiogenic drug, Charney et al., 1989; Morilak et al., 2005), and that ALLO (an anxiolytic neurosteroid, Patchev et al., 1996) did not attenuate yohimbine-induced responding in LoI rats, suggests that difference of treatment response was not due to heightened reactivity to an aversive event by LoI rats compared to HiI rats (theory discussed in chapter 4).

Rather, HiI rats may represent the abnormal treatment-resistant population. Studies of the effect of neurosteroids in humans have been mixed but present an interesting consideration. Progesterone, a precursor of ALLO (de Wit et al., 2001; Reddy, 2010), has been reported to decrease subjective "good" feeling associated with cocaine in human females that regularly used cocaine, thereby also decreasing motivation to obtain cocaine (Evans and Foltin, 2005). Progesterone decreased, to a lesser extent, the subjective good feeling in human male and females diagnosed as cocaine-dependent but did not affect self-administration of cocaine (Sofuoglu et al., 2004). It is unknown whether the smaller effect in Sofuoglu et al. (2004) was due to sex differences, as was found in the Evans and Foltin (2005) study, but in Reed et al. (2011), there was no effect on self-administration of cocaine or subjective effect of cocaine in human females. Thus, differences in receptivity to progesterone may have been due to differential degrees of dependence on cocaine, where the more dependent female subjects may not benefit from treatment, but the non-dependent cocaine users might. It should also be noted that subjects in the Evans and Foltin (2005) and Reed et al., (2011) studies may have been cocaine-dependent, but only in the Sofuoglu et al., (2004) study were subjects diagnosed as dependent.

Results from the study presented in this chapter indicate ALLO's effects on responding by LoI rats may be more representative of how treatment would affect average drug user, while HiI rats response to treatment may represent the abnormal user, or drug-dependent individual. Further support of this idea, as mentioned in the previous chapter, is found in high and low impulsive rats (selected with the 5-CSRT task). High

impulsive animals display abnormal behavior and abnormal underlying neurobiology in respect to corticostriatal control of cocaine-seeking behavior (Dalley et al., 2007; Murray et al., 2014; Chapter 3). In contrast, low impulsive rats' decrease in dopamine 2 receptor availability and transition of ventral to dorsal striatal control of behavior was more representative of the typical rat (Dalley et al., 2007; Murray et al., 2014; Chapter 4).

Similarly, LoI rats' response to ALLO was more typical of Wistar female rats (not selected or bred for high or low drug-seeking behaviors), whose cocaine-seeking responses were inhibited by ALLO (Anker and Carroll, 2010) and progesterone (Anker et al., 2007; Larson et al., 2007). The implications for the previous studies and the present result are that the cocaine-seeking responses in HiI rats may not have been due to positive reinforcement but rather driven by more automatic processes, such as habit-based decision-making or perhaps even more compulsive-type responses.

Thus, high drug-seeking animals may be more treatment resistant, whereas low drug-seeking animals tend to respond favorably to treatment. This idea is further supported by other research with humans, in which high discounting rates tended to predict poor treatment outcomes (Sheffer et al., 2012). Thus, treating impulsivity prior to treatment may help to reduce drug-using behaviors (discussed in the next chapter).

Improving treatment outcomes in treatment-resistant animals

Taken together, along with studies that indicate dorsolateral striatal (DLS) inactivation reduces reinstatement of drug-seeking behavior (See et al., 2007), these results and ideas suggest that DLS control of habit-based behavior (and perhaps, compulsivity) might be driving behavior more so in treatment-resistant animals. Prior studies have reported that

extinction of a habit requires new learning (Rescorla, 2001), and, as I have discussed in previous chapters, dorsomedial (DMS) striatum and corresponding pre-frontal/orbitofrontal cortical connections are central to flexible decision making and to the learning of new behaviors. Thus, effective treatments might include attempting to offset these more automatic processes by exploiting the flexible decision-making system or down regulating the habit-based decision-making system.

Previous reports have shown that inactivating DLS reduced habit-based behaviors by uncovering associations made during flexible decision making (Packard and McGaugh, 1996; Yin and Knowlton, 2004). Other studies have reported that inhibition of dopamine in the DMS reduced early-stage drug-seeking behavior but had no effect on late-stage drug-seeking behavior, whereas inhibiting dopamine in the DLS reduced drug seeking in late-stage drug seeking but not early-stage drug seeking (Murray et al., 2012). A similar effect was found in an pre-clinical alcohol study (Corbit et al., 2012a). Therefore, increasing activity in the DMS or decreasing activity in the DLS while revaluing the drug (e.g. devaluation, providing alternative rewards) may prove to be an effective means of reducing drug-seeking behaviors in high drug-seeking animals. This may improve goal-directed decision-making, allowing the agent to reassess the value of reward options.

At least in the case of the impulsive and saccharin-intake phenotypes, displaying differential high and low drug-seeking behavior, the high drug-seeking animals appear to have greater difficulty in controlling drug-seeking responses, even after punishment or prolonged extinction. This may indicate that increased habit-based behavior and/or

compulsion may have manifested itself more so in these animals. This is not to say that other phenotypes (e.g., novelty-seeking rats, Lewis and Fischer 344 rats, high and low cocaine responders) differ in drug use due to these same characteristics. For instance, in novelty-seeking rats, low novelty-responders tend to avoid aversive events (Dellu et al., 1996; Kabbaj et al., 2000; Stead et al., 2006) and may avoid drugs of abuse due to greater reactivity to the aversive side effects of drugs.

However, in animals whose choices are punishment-resistant, hence more habit-based, engaging flexible decision-making while devaluing the drug option or providing alternative rewards may draw choices away from drug consumption. This would have the effect of creating new associations and decreasing the probability of drug seeking. Contingency Management, a successful treatment for human drug addiction, appears to function in this manner.

As is the case with animal models of drug addiction, it is important to note that not all human drug addicts share this same characteristic. For example, as noted, individuals with mental health issues are more likely to be drug abusers. Thus, treating the mental health issue along with the drug addiction may be effective in minimizing drug-seeking behavior (Minkoff, 1989). We theorize, in the case of individuals with overactive habit-based decision-making systems, the behavioral treatment, Contingency Management, works extremely well, because it engages flexible decision-making systems (Chapter 6). In addition to a discussion of Contingency Management, I will discuss other factors that drive addictive behavior, successful treatment for these factors, and potentially new treatment for decreasing drug use for other factors.

Chapter 6: Translation to Human Addiction Treatment

In the previous chapters, the focus has been on addictive behavior in animals, potential mechanisms driving divergent drug-seeking behavior, potential treatment ideas, as well as potential biomarkers for treatment intervention. The goal of such research, in part, is to find applications for helping with the problem of human addiction.

As discussed in the previous chapter, animals that express overly-valued, habit-based, drug-seeking behavior tend to make more compulsive choices. This is likely due to control by a functionally overactive dorsolateral and motor/sensory corticostriatal system. Attempts to devalue actions controlled by this system have been ineffective (Adams, 1981, 1982), but inactivating dorsolateral striatum (DLS) renders actions amenable to devaluation again (Yin et al., 2004; Fuchs et al., 2006; See et al., 2007; Corbit et al., 2012; Jonkman et al., 2012; Gremel and Costa, 2013). Furthermore, as discussed, when the habit-based system is the primary control system, the dorsomedial and pre-frontal/orbitofrontal corticostriatal systems likely exhibit reduced functionality. Since inactivating whole brain regions in humans would be impractical and would cause significant problems even if affecting drug use, one way to counter the habit-based system is to engage the flexible-based system, either pharmacologically or with behavioral methods.

Our theory is that particular behavioral treatments are effective precisely because they are able to counter an overactive system by increasing activity in another competing system, thereby increasing the probability that problem drug use is attenuated. As an example, I will discuss one of these treatments, Contingency Management, which we

hypothesize is effective because it engages goal-directed decision making, countering more automatic processes, such as habit-based decision making.

Contingency Management

Contingency Management is one of the more effective methods of treating drug addiction, which it does by providing tangible rewards for proof of abstinence, affecting drug use. Studies in humans have shown that it is effective in reducing consumption of multiple types of drugs and increasing abstinence rates (Stitzer and Bigelow, 1978; Higgins et al., 1994; Petry et al., 2000, 2004, 2005; Carroll et al., 2002, 2006; Petry and Martin, 2002; Higgins et al., 2007). There are two main variations of Contingency Management, voucher-based (Higgins et al., 1994) and prized-based methods (Petry et al., 2000).

In the voucher-based method, drug-free urine samples are exchanged for points on a voucher that can be exchanged for consumer goods. Points are worth a specific amount of money, set by the researcher. The dollar value of points starts low but increases and accumulates over time as long as the individual remains abstinent. For example, in Higgins et al. (1994), points were worth $2.50 for the first clean sample and increased by $1.50 thereafter. By the end of the first month, if the individual remained abstinent, points were worth $16.50.

In the prize-based method, drug-free samples are exchanged for a chance to win a prize. Typically, prizes ranged from $1 to $100, with a high probability to win a small prize and a low probability to win medium and large prizes. In the Petry et al. (2000) study, the chance of winning anything worth over a dollar was less than 7%. A

comparison of the two methods found that, even though overall value of alternative rewards was lower in the prize-based method compared to the voucher-based method, the two methods were similarly effective (Petry et al., 2005)

Although the success of Contingency Management is thought to be due to the reinforcing effects of the alternative reward (points on a voucher or chance to win a prize), the pre-clinical data suggest that the value of the alternative reward would have to be much higher than those offered in Contingency Management to affect drug use. For example, studies have shown that the cost of the drug had to be increased by 100 (for low drug concentrations) to 1000 (for high drug concentrations) fold (Woolverton et al., 1997) in order to reduce drug consumption. Other studies have reported similar findings (Nader and Woolverton, 1991, 1992; Negus, 2003). Many of these studies implemented the use of only one lever to control the availability of the drug or the alternative reinforcer and another lever to deliver the reward. For instance, Woolverton and colleagues (1997) required monkeys to press one lever a certain number of times to change the cue light to green for cocaine and red for food. Subsequently, pressing the other lever would deliver either the cocaine or food. In this way, subjects were not required to use goal-directed behavior. Rather, it was likely more of a stimulus-response association (i.e., subjects pressed the lever until the cue-light changed to green, and the green cue light released a drug-seeking response).

In contrast, Ahmed and colleagues (Lenoir et al., 2007; Ahmed, 2010; Cantin et al., 2010) developed a choice procedure between a sweet liquid reward (typically a sucrose or saccharin solution) and cocaine by making cocaine available by pressing one

lever and liquid reward available by pressing the other lever. Interestingly, they reported significantly decreased cocaine choice in rats in the presence of an alternative reinforcer, the value of which was much smaller in magnitude than previous alternative reward studies (Lenoir et al., 2007). Similar results were found in a pre-clinical nicotine study (LeSage, 2009). They also tested the rats on a break-point procedure (measure of motivated reward-seeking behavior), and found that the same rats that had higher breakpoints for cocaine than the sweet liquid reward when cocaine or saccharin was presented alone preferred the sweet liquid reward when the two rewards were presented together (Cantin et al., 2010).

In all of these previous studies, what is clear is that the availability of an alternative reinforcer reduced drug-seeking behavior. However, it was the context in which the alternative reward was available that greatly affected probability that either the drug or the alternative reward was chosen. This suggests that rewards are not transituational (i.e., efficacy of reward changes across different experimental conditions (Meehl, 1950; Bickel and Madden, 1999)). In other words, a revealed preference condition, wherein a choice between two options is available, leads to a different valuation of reward when compared to a willingness to pay condition, in which the availability of reward is set against increasing cost.

Based on this information, we have proposed a new theory for why Contingency Management is effective (Regier and Redish, 2012, 2014). Alternative reward values in Contingency Management are low, especially when first starting treatment. We have

suggested that the value of rewards (especially in the first month) would not be enough to significantly reduce drug use.

We supported our theory by modeling the change in drug use as a function of cost using demand curves (Bruner and Johnson, 2014). In Contingency Management, cost is determined by the value of the alternative reinforcer, since increasing the value increases the opportunity cost of using drugs (using drugs means that the alternative reward is lost). Alternative reinforcement theory predicts that cost of the drug and decrease in use should be proportional. Conversely, our model found that the low value of alternative rewards at the beginning of treatment and the end of the first month of treatment would have produced a negligible effect on drug consumption (Figures 14 and 15).

We concluded that Contingency Management was inadequately explained by alternative reinforcement theory, that decision-making was not simply a basic cost-benefit analysis. Instead, we suggested that multiple decisions-making systems interact and compete to produce an action. We hypothesized that Contingency Management engages goal-directed decision-making by providing concrete, immediately available rewards for not using drugs, as compared to abstinence, which has more abstract and delayed rewards (Trope and Liberman, 2003; Heyman, 2009). The engagement of goal-directed decision-making systems competes with more automatic systems, such as the habit-based decision-making system, increasing the probability of subjects in Contingency Management treatment not using the drug option.

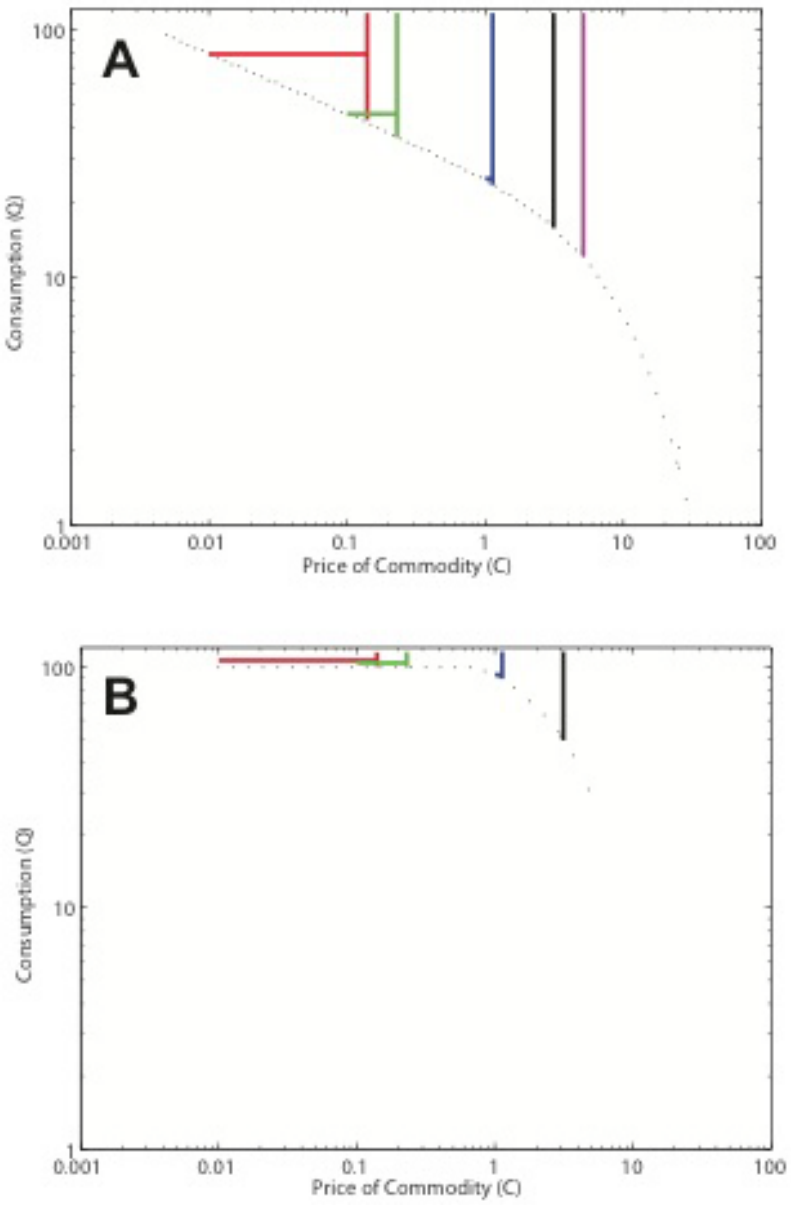


Figure 14. Expected consumption changes from a $2.50 voucher, given demand curves from Bruner and Johnson (2013). The five arrow-line pairs show predicted transitions produced by the increase in cost from losing the vouchers. a) Mean voucher curve averaged over all subjects. b) Example curve from a specific subject. Data replotted from Bruner and Johnson (2013) after translation with DataThief (www.datathief.org).

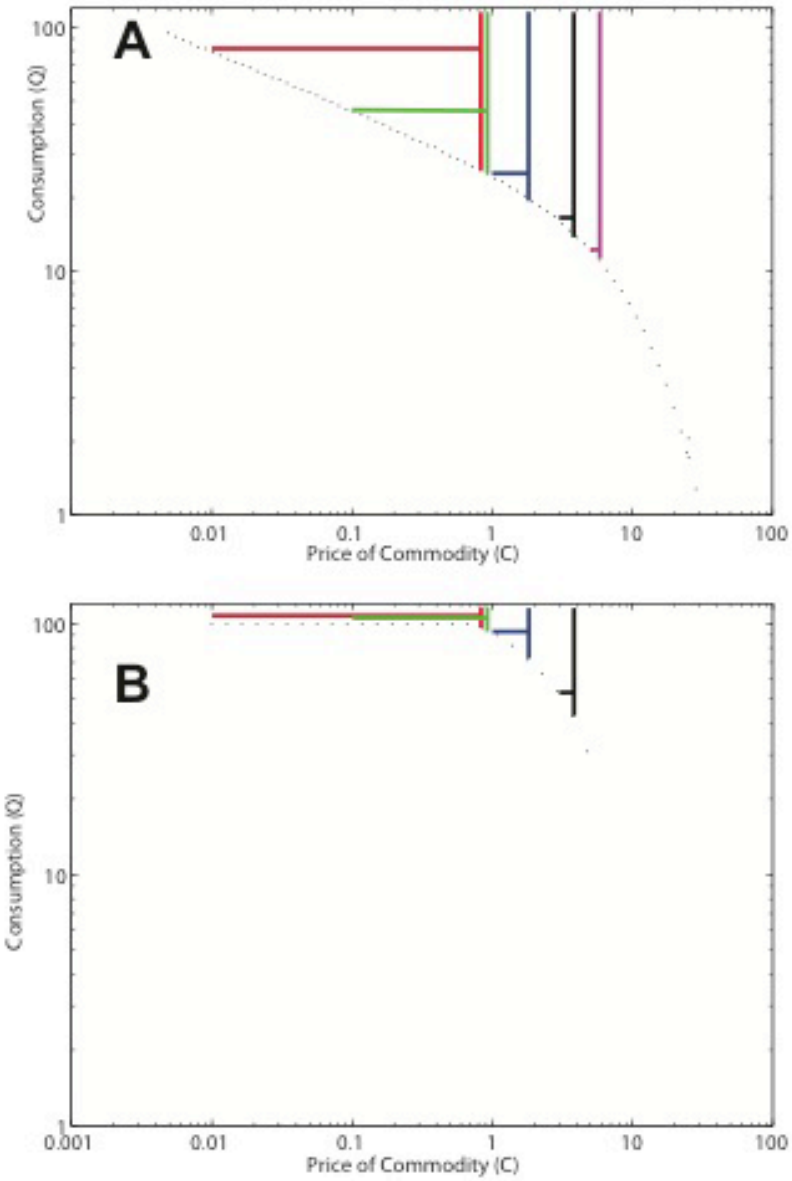


Figure 15. Expected consumption changes from a $16.25 voucher, given demand curves from Bruner and Johnson (2013). The five arrow-line pairs show predicted transitions produced by the increase in cost from losing the vouchers. (a) Mean voucher curve averaged over all subjects. (b) Example curve from a specific subject. Data replotted from Bruner and Johnson (2013) after translation with DataThief (www.datathief.org).

The predicted implications of our theory were that individuals who use drugs primarily driven by an overactive habit-based decision-making system and have robust flexible decision-making systems would benefit most from Contingency Management. We further predicted that individuals lacking robust flexible decision-making systems might first need to train the flexible decision-making system with behavioral or pharmacological interventions. Subsequently, after training, they may receive benefit from Contingency Management. Previous studies suggest that working memory training may be a promising candidate (Bickel et al., 2011).

Underlying neural correlates of this process would be similar to those discussed in the chapters of animal models of addiction. Habit-based behavior would be correlated with functional coupling of accumbens (analog of the rat ventral striatum) and putamen (analog of the rat dorsolateral striatum) along with reduced activity in the caudate (analog of the rat dorsomedial striatum). If our theories are correct, we would predict that differing activity levels of caudate (and pre-frontal/orbitofrontal cortical control) would predict varying levels of Contingency Management treatment success. Studies have shown that individuals with greater inhibitory control by the pre-frontal and orbitofrontal cortical areas of limbic systems prior to treatment predicted better treatment outcomes (Childress et al., 2015). What these studies have not shown are levels of pre-frontal or caudate activity in response to treatment. By showing that individuals who do not respond to Contingency Management also have attenuated corresponding activity in flexible decision-making systems, behavioral or pharamacological treatments could be

administered in an attempt to bolster flexible decision-making, and hence, response to Contingency Management.

Matching treatment with specific individual characteristics

Not everyone who responds poorly to Contingency Management would benefit from training up flexible decision-making systems, since, as discussed in the chapters on animal models of addition, it is likely individuals abuse drugs due to other mediating factors. As noted, some individuals may abuse drugs due to changes in homeostatic set points (Chapter 3), wherein stopping use of drugs after long-term use causes severe withdrawal. As noted previously, it is thought that a reduction of dopaminergic and opiate signaling and increased stress-inducing neuropeptides cause an allostatic state. This is a negative emotional state that changes valuation of potential choices, one in which the individual can choose to remain in a negative emotional state by not using drugs or alleviate the negative emotional state by using drugs, since using drugs brings these parameters back to normal, albeit temporarily. In this case, drug use would be driven by failures in the flexible decision-making system (Redish et al., 2008). Individuals in this allostatic state would engage in goal-directed behavior, weighing the options of remaining in that state or using drugs in an attempt to alleviate the negative emotional state. Treatment for this type of drug use has been to provide pharmacological treatments that reduce withdrawal, such as methadone maintenance (in the case of opiate addiction) treatment (Strain et al., 1994). Alternatively, the aim of some treatments have been to mitigate stress-causing peptides, such as corticotropin releasing factor, by

providing a corticotropin releasing factor antagonist (Koob, 2010), thus, reducing the motivation to alleviate a negative emotional state by using drugs.

Previous chapters have discussed impulsivity primarily in a pre-clinical context (Chapters 3, 4, and 5). The discussion is relevant to humans, as well, and is thought to represent another characteristic that can drive excessive drug use (de Wit and Richards, 2004). Theory suggests that more impulsive individuals tend to discount future gains more than non-impulsive individuals. This may cause an impulsive individual to value something immediately available (drug use) more so than something available further in the future (e.g. abstinence, saving money, career goals) (Kirby et al., 1999; Madden and Bickel, 2010). Thus, in the case of impulsivity, providing tools to help train cognitive systems to compete with impulsivity may be effective in reducing it, and subsequently, reducing drug use (Bickel et al., 2014). Two methods have already been mentioned. One is atomoxetine, shown to be effective in reducing impulsivity in animals (Economidou et al., 2011; Ansquer et al., 2014) and humans (Michelson et al., 2003; Faraone et al., 2005), and the other is working memory training, shown to reduce impulsivity in humans (Bickel et al., 2011b).

It has been suggested that divergent drug-seeking behavior in animal models of addiction may be driven by deficiencies in the dopamine system (Berridge and Robinson, 1998; Belin and Everitt, 2008; Everitt and Robbins, 2013). Similarly, variability in the dopamine system in humans has been suggested to drive differences in human drug use. For example, lower dopamine 2/3 receptors (D2R) and hypoactivity of the dopamine system has been correlated with human drug addiction, posited to cause decreased

sensitivity to drugs of abuse (Volkow et al., 2009). When those same individuals with hypoactive dopamine systems consumed drugs, the drugs caused hyperactivity in the dopamine system (Volkow et al., 2009). By restoring balance in the dopamine system (e.g., increasing D2R availability, compensatory dopamine interventions), drug use may be reduced (Volkow et al., 2003). However, manipulations to the dopamine system have been problematic, since changes to the dopamine system have tended to effect responsiveness to non-drug rewards, as well as salience to drug-associated stimuli, resulting in increased drug craving (Volkow, 2006).

Alternatively, addicts who are debilitated due to an imbalance in the dopamine systems may benefit from Contingency Management, since drug rewards are thought to be over-valued compared to non-drug rewards in these individuals. Increasing the value of non-drug options (by providing a concrete, immediately available reward for proof of abstinence, as is done in Contingency Management) might help to offset the overvaluation of drug rewards. Providing the optimal value of non-drug alternative rewards would be especially important. A meta-analysis of Contingency Management reported that higher initial value for the alternative reward in Contingency Management was more effective than lower initial values (Lussier et al., 2006). Perhaps, increased value of alternative rewards, resulting in increased effectiveness, was due to the treatment addressing two of the different type of the factors that drive addiction: an overactive habit-based system and a dopamine system imbalance.

Chapter 7: From rat neurophysiology to human behavior

In this dissertation, I have discussed how dorsal striatum is involved with decision-making systems and addiction. In particular, I have focused on individual variability in drug-seeking behaviors. In order to better understand addiction, I have integrated research from several different fields, including neurophysiology, immunohistochemistry, behavioral pharmacology, and theory. By integrating multiple perspectives, addiction can be understood from a broad perspective. Surveys report that addiction to illicit drugs and alcohol is a nationwide problem affecting nearly 10% of the population. Approximately 20% of the population engages in problem drinking; and nearly 10% of the population uses recreational drugs (SAHMSA, 2014). By understanding addiction from a broad perspective, several potential intervention points might be found, ranging from neurobiology to behavior.

I provided an overview of decision-making systems (Chapter 1) and the role of subsections of the dorsal striatum in these systems (Chapters 1 and 2). Flexible decision-making is typically employed during the learning of new environments (van der Meer et al., 2012b; Redish, 2013). The agent learns that specific cues, such as direction on a spatial task or a light on an operant task, predicts specific actions that lead to reward. Several regions, including the dorsomedial striatum (DMS) (Devan et al., 1999; Yin and Knowlton, 2004; Yin et al., 2005; Ragozzino, 2007) have been found to underlie this learning process. After extended experience in a predictable environment, actions are cached, and behavior becomes habit-based. The same stimuli (e.g., a left turn or a red

light) that previously predicted future outcomes during flexible decision making, release actions during habit-based decision making. The motor and sensory cortex and the dorsolateral striatum (DLS) have been found to underlie habit-based decision-making (Dickinson, 1985; Knowlton et al., 1996; Packard and McGaugh, 1996; Jog, 1999; Graybiel, 2008; Dezfouli and Balleine, 2013; Everitt, 2014; Chapters 1 and 2). Automatizing (or caching) stimulus-response actions in habit-based systems allows for other learning systems to make new associations. However, once stimulus-response actions are cached, they become difficult to change, even when the value of the reward changes (Adams and Dickinson, 1981; Adams, 1982). Thus, behaviors may persist even if disadvantageous.

Neurophysiological studies, have reported a development of DLS activity as behavior becomes more rigid, and that DLS drives this habit-based behavior (Hikosaka et al., 1995, 2002; Jog, 1999; Barnes et al., 2005; Thorn et al., 2010; van der Meer et al., 2010). The neurophysiological studies, thus far, for DMS have been limited. One study found that anterior DMS (aDMS) developed activity in the middle of a basic cued task (Thorn et al., 2010). Another study found that aDMS was not correlated with reversal learning (Kimchi and Laubach, 2009), and another study found that posterior DMS (pDMS) contains a small percentage of reward-prediction error neurons (Stalnaker et al., 2012). Lesions studies have shown that pDMS is more involved with flexible decision making than both DLS and aDMS, but no one has recorded from pDMS during a task that requires flexible behavior.

Thus, we investigated the neuronal activity of aDLS and pDMS on a spatial navigation task that required rats to change their behavioral strategy midway through test sessions. Our study (Chapter 2) found that as rats became more experienced on the task, within a single session, aDLS neurons increased firing at the maze location marking the end and beginning of each lap (as compared to the firing on the rest of the maze, a measure called the task-bracketing index (Smith and Graybiel, 2013), an effect observed before and after the switch. Posterior DMS neurons, on the other hand, did not develop task bracketing over laps. Surprisingly, though pDMS task bracketing tended to decrease over laps, it did not significantly decrease over laps as expected. Although, with aDLS task bracketing increasing and pDMS task bracketing slight decreasing, there was a significant difference of task-bracketing index between aDLS and pDMS on late laps.

Further, our results indicated a difference between aDLS and pDMS neurons in how they responded to the rats' position on the maze that was dependent on lap side and behavioral strategy. Neurons in the aDLS were more fixed to one side or the other. In contrast, pDMS neurons responded similarly to several points on the maze regardless of it being a left or right lap. Interestingly, firing of pDMS neurons changed before and after the switch in contingency. Thus, pDMS neurons were correlated with a change in behavioral strategy.

Therefore, our results indicated that pDMS neurons were important for initial learning and relearning of events in a changing environment. In contrast, aDLS neurons were important for the development of habit-based behaviors. Discussed in chapters 1, 2, and 3, as behavior changes from more flexible to more habit-based, neural control shifts

from ventral to dorsal striatum (Everitt and Robbins, 2013). Our data and recent studies suggest that DMS is important for learning. Thus, neural control may shift from medial (both ventral striatum and DMS) to DLS, perhaps even from ventral striatum to DMS to DLS. Actions become automatized over time, allowing for flexible systems to learn new action-outcome pairs (Chapters 1, 2, and 3). While this is a typical phenomenon with natural rewards, in addiction, this habit-based behavior can become maladaptive.

Some individuals are more susceptible to drug-seeking behaviors than others. Dysfunction in habit-based and flexible decision-making systems may be correlated with divergent drug-seeking behavior. Results from our study and previous studies (Chapter 4) indicate that high drug-seeking animals were more punishment-resistant than low drug-seeking animals. This could mean that high drug-seeking animals are more compulsive (Economidou et al., 2009; Holtz et al., 2013). Thus, hyperactive habit-based systems and hypoactive flexible systems may be an underlying factor in high drug-seeking animals. Since noncompensable dopamine may be a driving factor in creating compulsive behaviors (Redish, 2004), differences in dopamine processing may also underlie such behavior (Berridge and Robinson, 1998; Volkow et al., 2009). Interestingly, our results along with previous studies (Chapters 4 and 5) did not support the theory that low drug-seeking animals are more sensitive to aversive events, thought to be a type of protective factor to increased drug seeking (Carroll et al., 2009). Instead, variability in neurobiology may be a better predictor of divergent drug-seeking behaviors.

Our study found differences in cocaine-induced c-Fos induction in low drug-seeking animals compared to higher drug-seeking animals (Regier et al., 2012).

Surprisingly, low drug-seeking animals had greater cocaine-induced c-Fos expression in several brain regions. However, lower c-Fos expression has been associated with less availability of specific dopamine receptors (Ruskin and Marshall, 1994; Drago et al., 1996; Moratalla et al., 1996; Welter et al., 2007), and other high drug-seeking animals have reduced dopamine receptor availability (Flores et al., 1998; Dalley et al., 2007; Flagel et al., 2010). Thus, variability in dopamine receptor availability may be a contributing factor in driving the observed differences in c-Fos induction. Another possible explanation for low drug-seeking animals having greater cocaine-induced c-Fos expression, is that a blunted response to cocaine in high drug-seeking animals causes them to need more than low drug-seeking animals to achieve a similar drug-induced state. Thus, individual variability in drug-seeking behavior, as has been observed in the rodent models of drug addiction used in our study (Chapters 4 and 5), may be explained by variability in dopamine-signaling systems.

Successful treatment of addiction may very well depend on the ability to recognize individual differences and find effective treatments for each. Just as higher drug-seeking animals have tended to exhibit punishment-resistant behaviors, they have also been found to be more treatment-resistant (Holtz and Carroll, 2011; Anker et al., 2012). Our study (Chapter 5) found that low drug-seeking rats were more responsive to treatment than high drug-seeking rats. Specific treatments have been effective in dealing with specific characteristics that are predictive of increased drug use. For example, both atomoxetine (Michelson et al., 2003; Faraone et al., 2005; Ansquer et al., 2014) and working memory training (Bickel et al., 2011b) reduced impulsivity (predictive of higher

drug-seeking), and atomoxetine reduced drug seeking (Economidou et al., 2011). It has yet to be determined whether reducing impulsivity would reduce drug seeking. Overall, recognizing and treating individual characteristics may improve overall effectiveness of reducing drug-seeking behavior.

We hypothesized that Contingency Management, an effective treatment for human drug addiction, is effective because it treats individuals with specific characteristics (Chapter 6). We hypothesized that those with intact higher cognitive systems (e.g., PFC) but with hyperactive habit-based systems would benefit the most from Contingency Management, because we hypothesized that this treatment engages flexible decision-making systems. We posited that engaging flexible decision-making systems would help offset the hyperactive habit-based system.

Furthermore, we hypothesized that those dependent upon drugs for other reasons (e.g., hyperactive *stress* system) would receive minimal benefit from Contingency Management but might benefit from other treatment (e.g., pharmacological treatment of stress, cognitive behavioral therapy). In fact, there are several treatments currently available that appear to affect only a small percentage of addicts (e.g., Contingency Management, 12-Step programs, Inpatient treatment, Community Reinforcement Program). Perhaps the reason for generally low success rates in addiction treatment is that each treatment is effective only for a particular subgroup of addicts. Thus, matching individual variability with effective treatments may greatly improve overall treatment of drug addiction. This remains to be tested.

This work has focused on overly active habit-based behavior driving maladaptive drug-seeking decisions, with potential rectification by disrupting habit-based choices and engaging flexible decision making systems. Partly, this focus has been due to finding relevant potential roles for the dorsal striatum. Thus, an obvious limitation of this work is the lack of a thorough discussion of other characteristics that drive drug seeking and the corresponding neural correlates.

Understanding the mechanisms that drive differential drug-seeking behavior in animals and humans, and, subsequently, potential for treatment interventions, is an ongoing process and will continue to be an important area of drug addiction research. Nonetheless, this discourse has made clear a common theme. Notably, that flexible decision-making is central to associating value to drug rewards, and once cached by habit-based systems, the value of drugs is notoriously difficult to change. This can lead to rigid action selection that is characteristic of the behavior of many (but not all) individuals struggling with addiction. Thus, methods to increase flexible decision-making while facilitating non-drug using choices may be helpful for many (but not all) individuals with addiction. In all, decision-making systems theory may provide a unifying framework for addiction (Redish et al., 2008), where different theories of addiction would be more relevant for specific individuals. Furthermore, individual variability would involve dysfunction in different decision-making systems with corresponding underlying neural correlates, implicating that targeting specific behaviors and neural correlates of behavior would improve treatment effectiveness.

Bibliography

Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q J Exp Psychol-B* 34: 77–98, 1982.

Adams CD, Dickinson A. Instrumental responding following reinforcer devaluation. *Q J Exp Psychol-B* 33: 109–121, 1981.

Ahmed SH. Validation crisis in animal models of drug addiction: beyond non-disordered drug use toward drug addiction. *Neurosci Biobehav Rev* 35: 172–184, 2010.

Ahmed SH, Koob GF. Transition from moderate to excessive drug intake: change in hedonic set point. *Science* 282: 298–300, 1998.

Alexander GE, Crutcher MD. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci* 13: 266–271, 1990.

Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 9: 357–381, 1986.

Anderson KG, Velkey AJ, Woolverton WL. The generalized matching law as a predictor of choice between cocaine and food in rhesus monkeys. *Psychopharmacology (Berl)* 163: 319–326, 2002.

Anker JJ, Carroll ME. Sex differences in the effects of allopregnanolone on yohimbine-induced reinstatement of cocaine seeking in rats. *Drug Alcohol Depend* 107: 264–267, 2010.

Anker JJ, Holtz NA, Carroll ME. Effects of progesterone on escalation of intravenous cocaine self-administration in rats selectively bred for high or low saccharin intake. *Behav Pharmacol* 23: 205–210, 2012.

Anker JJ, Perry JL, Gliddon LA, Carroll ME. Impulsivity predicts the escalation of cocaine self-administration in rats. *Pharmacol Biochem Behav* 93: 343–348, 2009.

Ansquer S, Belin-Rauscent A, Dugast E, Duran T, Benatru I, Mar AC, Houeto J-L, Belin D. Atomoxetine decreases vulnerability to develop compulsivity in high impulsive rats. *Biol Psychiatry* 75: 825–832, 2014.

Anthony JC, Warner LA, Kessler RC. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: Basic findings from the National Comorbidity Survey. *Exp Clin Psychopharmacol* 2: 244–268, 1994.

Atallah HE, Lopez-Paniagua D, Rudy JW, O'Reilly RC. Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat Neurosci* 10: 126–131, 2007.

Bacharach S, Kawa A, Caroll M, Calu D. Sign- and goal-tracking rats learn differently in the face of changing reward value. Society for Neuroscience. Washington D.C.: 2014.

Badia-Elder N, Kiefer SW, Dess NK. Taste reactivity in rats selectively bred for high vs. low saccharin consumption. *Physiol Behav* 59: 749–755, 1996.

Balleine BW. Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. *Physiol Behav* 86: 717–730, 2005.

Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437: 1158–1161, 2005.

Barto AG. *Reinforcement Learning: An Introduction*. MIT Press, 1998.

Belin D, Everitt BJ. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron* 57: 432–441, 2008.

Berendse HW, Galis-de Graaf Y, Groenewegen HJ. Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *J Comp Neurol* 316: 314–347, 1992.

Berke JD, Breck JT, Eichenbaum H. Striatal versus hippocampal representations during win-stay maze performance. *J Neurophysiol* 101: 1575–1587, 2009.

Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev* 28: 309–369, 1998.

Besson M, Pelloux Y, Dilleen R, Theobald DE, Lyon A, Belin-Rauscent A, Robbins TW, Dalley JW, Everitt BJ, Belin D. Cocaine modulation of frontostriatal expression of Zif268, D2, and 5-HT2c receptors in high and low impulsive rats. *Neuropsychopharmacology* 38: 1963–1973, 2013.

Bickel WK, Madden GJ. A comparison of measures of relative reinforcing efficacy and behavioral economics: cigarettes and money in smokers. *Behav Pharmacol* 10: 627–637, 1999.

Bickel WK, Miller ML, Yi R, Kowal BP, Lindquist DM, Pitcock JA. Behavioral and neuroeconomics of drug addiction: competing neural systems and temporal discounting processes. *Drug Alcohol Depend* 90 Suppl 1: S85–91, 2007.

Bickel WK, Yi R, Landes RD, Hill PF, Baxter C. Remember the future: working memory training decreases delay discounting among stimulant addicts. *Biol Psychiatry* 69: 260–265, 2011.

Blumenthal A, Steiner A, Seeland K, Redish AD. Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task. *Neurobiol Learn Mem* 95: 376–384, 2011.

Boulougouris V, Dalley JW, Robbins TW. Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behav Brain Res* 179: 219–228, 2007.

Bruner NR, Johnson MW. Demand curves for hypothetical cocaine in cocaine-dependent individuals. *Psychopharmacology (Berl)* 231: 889–897, 2014.

Buckner RL, Carroll DC. Self-projection and the brain. *Trends Cogn Sci* 11: 49–57, 2007.

Cantin L, Lenoir M, Augier E, Vanhille N, Dubreucq S, Serre F, Vouillac C, Ahmed SH. Cocaine is low on the value ladder of rats: possible evidence for resilience to addiction. *PLoS ONE* 5: e11592, 2010.

Capper-Loup C, Canales JJ, Kadaba N, Graybiel AM. Concurrent activation of dopamine D1 and D2 receptors is required to evoke neural and behavioral phenotypes of cocaine sensitization. *J Neurosci* 22: 6218–6227, 2002.

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

Carelli RM, West MO. Representation of the body by single neurons in the dorsolateral striatum of the awake, unrestrained rat. *J Comp Neurol* 309: 231–249, 1991.

Carroll KM, Easton CJ, Nich C, Hunkele KA, Neavins TM, Sinha R, Ford HL, Vitolo SA, Doebrick CA, Rounsaville BJ. The use of contingency management and motivational/skills-building therapy to treat young adults with marijuana dependence. *J Consult Clin Psychol* 74: 955–966, 2006.

Carroll KM, Sinha R, Nich C, Babuscio T, Rounsaville BJ. Contingency management to enhance naltrexone treatment of opioid dependence: a randomized clinical trial of reinforcement magnitude. *Exp Clin Psychopharmacol* 10: 54–63, 2002.

Carroll ME, Anderson MM, Morgan AD. Higher locomotor response to cocaine in female (vs. male) rats selectively bred for high (HiS) and low (LoS) saccharin intake. *Pharmacol Biochem Behav* 88: 94–104, 2007.

Carroll ME, Anker JJ, Perry JL. Modeling risk factors for nicotine and other drug abuse in the preclinical laboratory. *Drug Alcohol Depend* 104 Suppl 1: S70–78, 2009.

Carroll ME, Morgan AD, Anker JJ, Perry JL, Dess NK. Selective breeding for differential saccharin intake as an animal model of drug abuse. *Behav Pharmacol* 19: 435–460, 2008.

Castañé A, Theobald DEH, Robbins TW. Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behav Brain Res* 210: 74–83, 2010.

Charney DS, Woods SW, Heninger GR. Noradrenergic function in generalized anxiety disorder: effects of yohimbine in healthy subjects and patients with generalized anxiety disorder. *Psychiatry Res* 27: 173–182, 1989.

Chen BT, Yau H-J, Hatch C, Kusumoto-Yoshida I, Cho SL, Hopf FW, Bonci A. Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. *Nature* 496: 359–362, 2013.

Childress AR, Suh J, Jagannathan K, Young KA, Franklin T, Ehrman R, Monge Z, Langleben D, Kampman KM, Wetherill R, Gawrysiak MJ, O'Brien CP. "Destined for success"? Brain responses during drug cue exposure foretell clinical outcome in cocaine patients. *Drug & Alcohol Dependence* 146: e246, 2015.

Cho J, West MO. Distributions of single neurons related to body parts in the lateral striatum of the rat. *Brain Res* 756: 241–246, 1997.

Chudasama Y, Passetti F, Rhodes SEV, Lopian D, Desai A, Robbins TW. Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: differential effects on selectivity, impulsivity and compulsivity. *Behav Brain Res* 146: 105–119, 2003.

Cicero TJ, Nock B, Meyer ER. Gender-linked differences in the expression of physical dependence in the rat. *Pharmacol Biochem Behav* 72: 691–697, 2002.

Clarke HF, Robbins TW, Roberts AC. Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J Neurosci* 28: 10972–10982, 2008.

Corbit LH, Janak PH. Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *Eur J Neurosci* 31: 1312–1321, 2010.

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

Crider A, Solomon PR, McMahon MA. Disruption of selective attention in the rat following chronic d-amphetamine administration: relationship to schizophrenic attention disorder. *Biol Psychiatry* 17: 351–361, 1982.

Dackis CA, O'Brien CP. Cocaine dependence: a disease of the brain's reward centers. *J Subst Abuse Treat* 21: 111–117, 2001.

Dalley JW, Fryer TD, Brichard L, Robinson ESJ, Theobald DEH, Lääne K, Peña Y, Murphy ER, Shah Y, Probst K, Abakumova I, Aigbirhio FI, Richards HK, Hong Y, Baron J-C, Everitt BJ, Robbins TW. Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science* 315: 1267–1270, 2007a.

Dalley JW, Lääne K, Theobald DEH, Peña Y, Bruce CC, Huszar AC, Wojcieszek M, Everitt BJ, Robbins TW. Enduring deficits in sustained visual attention during withdrawal of intravenous methylenedioxymethamphetamine self-administration in rats: results from a comparative study with d-amphetamine and methamphetamine. *Neuropsychopharmacology* 32: 1195–1206, 2007b.

Davis BA, Clinton SM, Akil H, Becker JB. The effects of novelty-seeking phenotypes and sex differences on acquisition of cocaine self-administration in selectively bred High-Responder and Low-Responder rats. *Pharmacol Biochem Behav* 90: 331–338, 2008.

Daw N, Niv Y, Dayan P. Actions, policies, values, and the basal ganglia. In: *Recent Breakthroughs in Basal Ganglia Research*, edited by Bezard E. New York: Nova Science Publishers, 2006, p. 111–130.

Dayan P, Niv Y, Seymour B, Daw ND. The misbehavior of value and the discipline of the will. *Neural Netw* 19: 1153–1160, 2006.

Dellu F, Piazza PV, Mayo W, Le Moal M, Simon H. Novelty-seeking in rats--biobehavioral characteristics and possible relationship with the sensation-seeking trait in man. *Neuropsychobiology* 34: 136–145, 1996.

Deroche-Gamonet V, Belin D, Piazza PV. Evidence for addiction-like behavior in the rat. *Science* 305: 1014–1017, 2004.

Dess N, Arnal J, Chapman C, Siebel S, VenderWeele D, Green K. Exploring adaptations to famine: rats selectively bred for differential saccharin intake differ on deprivation-induced hyperactivity and emotionality. *Int J Comp Psychol* 13: 34–53, 2000.

Dess NK. Responses to basic taste qualities in rats selectively bred for high versus low saccharin intake. *Physiol Behav* 69: 247–257, 2000.

Dess N, Minor T. Taste and emotionality in rats selectively bred for high versus low saccharin intake. *Anim Learn Behav* 24: 105–115, 1996.

Devan BD, McDonald RJ, White NM. Effects of medial and lateral caudate-putamen lesions on place- and cue-guided behaviors in the water maze: relation to thigmotaxis. *Behav Brain Res* 100: 5–14, 1999.

Dezfouli A, Balleine BW. Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Comput Biol* 9: e1003364, 2013.

Dickinson A. Actions and habits: The development of behavioural autonomy. *Philos T Roy Soc B* 308: 67–78, 1985.

Dietz DM, Tapocik J, Gaval-Cruz M, Kabbaj M. Dopamine transporter, but not tyrosine hydroxylase, may be implicated in determining individual differences in behavioral sensitization to amphetamine. *Physiol Behav* 86: 347–355, 2005.

Drago J, Gerfen CR, Westphal H, Steiner H. D1 dopamine receptor-deficient mouse: cocaine-induced regulation of immediate-early gene and substance P expression in the striatum. *Neuroscience* 74: 813–823, 1996.

Economidou D, Dalley JW, Everitt BJ. Selective norepinephrine reuptake inhibition by atomoxetine prevents cue-induced heroin and cocaine seeking. *Biol Psychiatry* 69: 266–274, 2011.

Economidou D, Pelloux Y, Robbins TW, Dalley JW, Everitt BJ. High impulsivity predicts relapse to cocaine-seeking after punishment-induced abstinence. *Biol Psychiatry* 65: 851–856, 2009.

Erb S, Shaham Y, Stewart J. Stress reinstates cocaine-seeking behavior after prolonged extinction and a drug-free period. *Psychopharmacology (Berl)* 128: 408–412, 1996.

Ersche KD, Roiser JP, Abbott S, Craig KJ, Müller U, Suckling J, Ooi C, Shabbir SS, Clark L, Sahakian BJ, Fineberg NA, Merlo-Pich EV, Robbins TW, Bullmore ET. Response perseveration in stimulant dependence is associated with striatal

dysfunction and can be ameliorated by a D(2/3) receptor agonist. *Biol Psychiatry* 70: 754–762, 2011.

Eshel N, Roiser J. Reward and punishment processing in depression. *Biol Psychiat* 68: 118–124, 2010.

Evans SM, Foltin RW. Exogenous progesterone attenuates the subjective effects of smoked cocaine in women, but not in men. *Neuropsychopharmacology* 31: 659–674, 2006.

Everitt BJ. Neural and psychological mechanisms underlying compulsive drug seeking habits and drug memories--indications for novel treatments of addiction. *Eur J Neurosci* 40: 2163–2182, 2014.

Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci* 8: 1481–1489, 2005.

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

Faraone SV, Biederman J, Spencer T, Michelson D, Adler L, Reimherr F, Glatt SJ. Efficacy of atomoxetine in adult attention-deficit/hyperactivity disorder: a drug-placebo response curve analysis. *Behav Brain Funct* 1: 16, 2005.

Feltenstein MW, See RE. Potentiation of cue-induced reinstatement of cocaine-seeking in rats by the anxiogenic drug yohimbine. *Behav Brain Res* 174: 1–8, 2006.

Ferry AT, Lu XC, Price JL. 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, 2000.

Flagel SB, Akil H, Robinson TE. Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology* 56 Suppl 1: 139–148, 2009.

Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PEM, Akil H. A selective role for dopamine in stimulus-reward learning. *Nature* 469: 53–57, 2011.

Flagel SB, Robinson TE, Clark JJ, Clinton SM, Watson SJ, Seeman P, Phillips PEM, Akil H. An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: implications for addiction. *Neuropsychopharmacology* 35: 388–400, 2010.

Flores G, Wood GK, Barbeau D, Quirion R, Srivastava LK. Lewis and Fischer rats: a comparison of dopamine transporter and receptors levels. *Brain Res* 814: 34–40, 1998.

Fuchs RA, Branham RK, See RE. Different neural substrates mediate cocaine seeking after abstinence versus extinction training: a critical role for the dorsolateral caudate-putamen. *J Neurosci* 26: 3584–3588, 2006.

Gabriele A, See RE. Reversible inactivation of the basolateral amygdala, but not the dorsolateral caudate putamen, attenuates consolidation of cocaine-cue associative learning in a reinstatement model of drug-seeking. *Eur J Neurosci* 32: 1024–1029, 2010.

Glowa JR, Shaw AE, Riley AL. Cocaine-induced conditioned taste aversions: comparisons between effects in LEW/N and F344/N rat strains. *Psychopharmacology (Berl)* 114: 229–232, 1994.

Grabus SD, Glowa JR, Riley AL. Morphine- and cocaine-induced c-Fos levels in Lewis and Fischer rat strains. *Brain Res* 998: 20–28, 2004.

Graybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem* 70: 119–136, 1998.

Graybiel AM. The basal ganglia: learning new tricks and loving it. *Curr Opin Neurobiol* 15: 638–644, 2005.

Graybiel AM. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31: 359–387, 2008.

Graybiel AM, Moratalla R, Robertson HA. Amphetamine and cocaine induce drug-specific activation of the c-fos gene in striosome-matrix compartments and limbic subdivisions of the striatum. *Proc Natl Acad Sci USA* 87: 6912–6916, 1990.

Graybiel AM, Pickel VM, Joh TH, Reis DJ, Ragsdale CW. Direct demonstration of a correspondence between the dopamine islands and acetylcholinesterase patches in the developing striatum. *Proc Natl Acad Sci USA* 78: 5871–5875, 1981.

Graybiel AM, Rauch SL. Toward a neurobiology of obsessive-compulsive disorder. *Neuron* 28: 343–347, 2000.

Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* 4: 2264, 2013.

Hebb D, Williams K. A method of rating animal intelligence. *J Gen Psychol* 34: 59–65, 1946.

Heishman SJ, Henningfield JE. Tolerance to repeated nicotine administration on performance, subjective, and physiological responses in nonsmokers. *Psychopharmacology (Berl)* 152: 321–333, 2000.

Heyman GM. *Addiction: A Disorder of Choice*. Harvard University Press, 2009.

Higgins ST, Budney AJ, Bickel WK, Foerg FE, Donham R, Badger GJ. Incentives improve outcome in outpatient behavioral treatment of cocaine dependence. *Arch Gen Psychiatry* 51: 568–576, 1994.

Hikosaka O, Nakamura K, Sakai K, Nakahara H. Central mechanisms of motor skill learning. *Curr Opin Neurobiol* 12: 217–222, 2002.

Hikosaka O, Rand MK, Miyachi S, Miyashita K. Learning of sequential movements in the monkey: process of learning and retention of memory. *J Neurophysiol* 74: 1652–1661, 1995.

Hogle JM, Curtin JJ. Sex differences in negative affective response during nicotine withdrawal. *Psychophysiology* 43: 344–356, 2006.

Holtz NA, Carroll ME. Baclofen has opposite effects on escalation of cocaine self-administration: increased intake in rats selectively bred for high (HiS) saccharin intake and decreased intake in those selected for low (LoS) saccharin intake. *Pharmacol Biochem Behav* 100: 275–283, 2011.

Holtz NA, Anker JJ, Regier PS, Claxton A, Carroll ME. Cocaine self-administration punished by i.v. histamine in rat models of high and low drug abuse vulnerability: effects of saccharin preference, impulsivity, and sex. *Physiol Behav* 122: 32–38, 2013.

Hooks MS, Jones GH, Smith AD, Neill DB, Justice JB. Individual differences in locomotor activity and sensitization. *Pharmacol Biochem Behav* 38: 467–470, 1991.

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

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

Jaffe, Gitisetan, Tarash, Pham, Jentsch. Are nicotine-related cues susceptible to the blocking effect? Society for Neuroscience. 2010.

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. Building neural representations of habits. *Science* 286: 1745–1749, 1999.

Johnson A, Redish AD. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J Neurosci* 27: 12176–12189, 2007.

Kabbaj M, Akil H. Individual differences in novelty-seeking behavior in rats: a c-fos study. *Neuroscience* 106: 535–545, 2001.

Kabbaj M, Devine DP, Savage VR, Akil H. Neurobiological correlates of individual differences in novelty-seeking behavior in the rat: differential expression of stress-related molecules. *J Neurosci* 20: 6983–6988, 2000.

Kamin L, Campbell B (ed), Church B (ed). Predictability, surprise, attention, and conditioning. In: *Punishment and Aversive Behaviour*. New-York: Appleton-Century-Croft, 1969, p. pp 279– 296.

Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* 13: 400–408, 2003.

Kimchi EY, Laubach M. The dorsomedial striatum reflects response bias during learning. *J Neurosci* 29: 14891–14902, 2009.

Kimura M. Behaviorally contingent property of movement-related activity of the primate putamen. *J Neurophysiol* 63: 1277–1296, 1990.

Kirby KN, Petry NM, Bickel WK. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J Exp Psychol Gen* 128: 78–87, 1999.

Kirkby R. Caudate nucleus lesions and perseverative behavior. *Physiol Behav* 4: 451–454, 1969.

Knowlton BJ, Mangels JA, Squire LR. A neostriatal habit learning system in humans. *Science* 273: 1399–1402, 1996.

Kolb B. Studies on the caudate-putamen and the dorsomedial thalamic nucleus of the rat: implications for mammalian frontal-lobe functions. *Physiol Behav* 18: 237–244, 1977.

Konradi C, Leveque JC, Hyman SE. Amphetamine and dopamine-induced immediate early gene expression in striatal neurons depends on postsynaptic NMDA receptors and calcium. *J Neurosci* 16: 4231–4239, 1996.

Koob GF. The neurobiology of addiction: a neuroadaptational view relevant for diagnosis. *Addiction* 101 Suppl 1: 23–30, 2006.

Koob GF. The role of CRF and CRF-related peptides in the dark side of addiction. *Brain Res* 1314: 3–14, 2010.

Koob GF, Le Moal M. Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* 24: 97–129, 2001.

Koob GF, Sanna PP, Bloom FE. Neuroscience of addiction. *Neuron* 21: 467–476, 1998.

Kosten TA, Ambrosio E. HPA axis function and drug addictive behaviors: insights from studies with Lewis and Fischer 344 inbred rats. *Psychoneuroendocrinology* 27: 35–69, 2002.

Kosten TA, Miserendino MJ, Chi S, Nestler EJ. Fischer and Lewis rat strains show differential cocaine effects in conditioned place preference and behavioral sensitization but not in locomotor activity or conditioned taste aversion. *J Pharmacol Exp Ther* 269: 137–144, 1994.

Kosten TA, Miserendino MJ, Haile CN, DeCaprio JL, Jatlow PI, Nestler EJ. Acquisition and maintenance of intravenous cocaine self-administration in Lewis and Fischer inbred rat strains. *Brain Res* 778: 418–429, 1997.

Larson EB, Anker JJ, Gliddon LA, Fons KS, Carroll ME. Effects of estrogen and progesterone on the escalation of cocaine self-administration in female rats during extended access. *Exp Clin Psychopharmacol* 15: 461–471, 2007.

Lenoir M, Serre F, Cantin L, Ahmed SH. Intense sweetness surpasses cocaine reward. *PLoS ONE* 2: e698, 2007.

Lesage MG. Toward a nonhuman model of contingency management: effects of reinforcing abstinence from nicotine self-administration in rats with an alternative nondrug reinforcer. *Psychopharmacology (Berl)* 203: 13–22, 2009.

Leshner AI. Addiction is a brain disease, and it matters. *Science* 278: 45–47, 1997.

Letchworth SR, Nader MA, Smith HR, Friedman DP, Porrino LJ. Progression of changes in dopamine transporter binding site density as a result of cocaine self-administration in rhesus monkeys. *J Neurosci* 21: 2799–2807, 2001.

Lex B, Hauber W. Disconnection of the entorhinal cortex and dorsomedial striatum impairs the sensitivity to instrumental contingency degradation. *Neuropsychopharmacology* 35: 1788–1796, 2010.

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

Lussier JP, Heil SH, Mongeon JA, Badger GJ, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction* 101: 192–203, 2006.

Marks KR, Kearns DN, Christensen CJ, Silberberg A, Weiss SJ. Learning that a cocaine reward is smaller than expected: A test of Redish's computational model of addiction. *Behav Brain Res* 212: 204–207, 2010.

McGeorge AJ, Faull RL. The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* 29: 503–537, 1989.

Meehl PE. On the circularity of the law of effect. *Psychological Bulletin* 47: 52–75, 1950.

Van der Meer MAA, Johnson A, Schmitzer-Torbert NC, Redish AD. Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron* 67: 25–32, 2010.

Van der Meer M, Kurth-Nelson Z, Redish AD. Information processing in decision-making systems. *The Neuroscientist* 18: 342–359, 2012.

Menzies L, Chamberlain SR, Laird AR, Thelen SM, Sahakian BJ, Bullmore ET. Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: the orbitofronto-striatal model revisited. *Neurosci Biobehav Rev* 32: 525–549, 2008.

Michelson D, Adler L, Spencer T, Reimherr FW, West SA, Allen AJ, Kelsey D, Wernicke J, Dietrich A, Milton D. Atomoxetine in adults with ADHD: two randomized, placebo-controlled studies. *Biol Psychiatry* 53: 112–120, 2003.

Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. 1956. *Psychol Rev* 101: 343–352, 1994.

Minkoff K. An integrated treatment model for dual diagnosis of psychosis and addiction. *Hosp Community Psychiatry* 40: 1031–1036, 1989.

Miyachi S, Hikosaka O, Lu X. Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Exp Brain Res* 146: 122–126, 2002.

Miyachi S, Hikosaka O, Miyashita K, Kárádi Z, Rand MK. Differential roles of monkey striatum in learning of sequential hand movement. *Exp Brain Res* 115: 1–5, 1997.

Molander AC, Mar A, Norbury A, Steventon S, Moreno M, Caprioli D, Theobald DEH, Belin D, Everitt BJ, Robbins TW, Dalley JW. High impulsivity predicting vulnerability to cocaine addiction in rats: some relationship with novelty preference but not novelty reactivity, anxiety or stress. *Psychopharmacology (Berl)* 215: 721–731, 2011.

Moore RJ, Vinsant SL, Nader MA, Porrino LJ, Friedman DP. Effect of cocaine self-administration on dopamine D2 receptors in rhesus monkeys. *Synapse* 30: 88–96, 1998.

Moratalla R, Xu M, Tonegawa S, Graybiel AM. Cellular responses to psychomotor stimulant and neuroleptic drugs are abnormal in mice lacking the D1 dopamine receptor. *Proc Natl Acad Sci USA* 93: 14928–14933, 1996.

Muenzinger K. Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *The Pedagogical Seminary and Journal of Genetic Psychology* 53: 75–86, 1938.

Murray JE, Belin D, Everitt BJ. Double dissociation of the dorsomedial and dorsolateral striatal control over the acquisition and performance of cocaine seeking. *Neuropsychopharmacology* 37: 2456–2466, 2012.

Murray JE, Dilleen R, Pelloux Y, Economidou D, Dalley JW, Belin D, Everitt BJ. Increased impulsivity retards the transition to dorsolateral striatal dopamine control of cocaine seeking. *Biol Psychiatry* 76: 15–22, 2014.

Nader MA, Daunais JB, Moore T, Nader SH, Moore RJ, Smith HR, Friedman DP, Porrino LJ. Effects of cocaine self-administration on striatal dopamine systems in rhesus monkeys: initial and chronic exposure. *Neuropsychopharmacology* 27: 35–46, 2002.

Nader MA, Hedeker D, Woolverton WL. Behavioral economics and drug choice: effects of unit price on cocaine self-administration by monkeys. *Drug Alcohol Depend* 33: 193–199, 1993.

Nader MA, Woolverton WL. Effects of increasing response requirement on choice between cocaine and food in rhesus monkeys. *Psychopharmacology (Berl)* 108: 295–300, 1992.

Negus SS. Rapid assessment of choice between cocaine and food in rhesus monkeys: effects of environmental manipulations and treatment with d-amphetamine and flupenthixol. *Neuropsychopharmacology* 28: 919–931, 2003.

Nelson A, Killcross S. Amphetamine exposure enhances habit formation. *J Neurosci* 26: 3805–3812, 2006.

Nelson AM, Kleschen MJ, Zahniser NR. Individual differences in cocaine-induced locomotor activity of male Sprague-Dawley rats are not explained by plasma corticosterone levels. *Neurosci Lett* 476: 9–13, 2010.

Nestler EJ. Molecular neurobiology of addiction. *The American Journal on Addictions* 10: 201–217, 2001.

Nestler EJ. Historical review: Molecular and cellular mechanisms of opiate and cocaine addiction. *Trends Pharmacol Sci* 25: 210–218, 2004.

Niv Y, Joel D, Dayan P. A normative perspective on motivation. *Trends Cogn Sci* 10: 375–381, 2006.

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

Ohad D, Lubow RE, Weiner I, Feldon J. The effects of amphetamine on blocking. *Psychobiology* 15: 137–143, 1987.

Olmstead MC, Lafond MV, Everitt BJ, Dickinson A. Cocaine seeking by rats is a goal-directed action. *Behav Neurosci* 115: 394–402, 2001.

O'Tuathaigh CMP, Salum C, Young AMJ, Pickering AD, Joseph MH, Moran PM. The effect of amphetamine on Kamin blocking and overshadowing. *Behav Pharmacol* 14: 315–322, 2003.

Packard MG, Hirsh R, White NM. Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J Neurosci* 9: 1465–1472, 1989.

Packard MG, McGaugh JL. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Mem* 65: 65–72, 1996.

Panlilio LV, Thorndike EB, Schindler CW. Blocking of conditioning to a cocaine-paired stimulus: testing the hypothesis that cocaine perpetually produces a signal of larger-than-expected reward. *Pharmacol Biochem Behav* 86: 774–777, 2007.

Papale AE, Stott JJ, Powell NJ, Regier PS, Redish AD. Interactions between deliberation and delay-discounting in rats. *Cogn Affect Behav Neurosci* 12: 513–526, 2012.

Parent A. *Comparative neurobiology of the basal ganglia*. J. Wiley, 1986.

Parent A, Hazrati LN. Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain Res Brain Res Rev* 20: 91–127, 1995.

Parent A, Sato F, Wu Y, Gauthier J, Lévesque M, Parent M. Organization of the basal ganglia: the importance of axonal collateralization. *Trends Neurosci* 23: S20–27, 2000.

Patchev VK, Hassan AH, Holsboer DF, Almeida OF. The neurosteroid tetrahydroprogesterone attenuates the endocrine response to stress and exerts glucocorticoid-like effects on vasopressin gene transcription in the rat hypothalamus. *Neuropsychopharmacology* 15: 533–540, 1996.

Pelloux Y, Everitt BJ, Dickinson A. Compulsive drug seeking by rats under punishment: effects of drug taking history. *Psychopharmacology (Berl)* 194: 127–137, 2007.

Perry JL, Anderson MM, Nelson SE, Carroll ME. Acquisition of i.v. cocaine self-administration in adolescent and adult male rats selectively bred for high and low saccharin intake. *Physiol Behav* 91: 126–133, 2007.

Perry JL, Larson EB, German JP, Madden GJ, Carroll ME. Impulsivity (delay discounting) as a predictor of acquisition of IV cocaine self-administration in female rats. *Psychopharmacology (Berl)* 178: 193–201, 2005.

Perry JL, Morgan AD, Anker JJ, Dess NK, Carroll ME. Escalation of i.v. cocaine self-administration and reinstatement of cocaine-seeking behavior in rats bred for high and low saccharin intake. *Psychopharmacology (Berl)* 186: 235–245, 2006.

Perry JL, Nelson SE, Carroll ME. Impulsive choice as a predictor of acquisition of IV cocaine self- administration and reinstatement of cocaine-seeking behavior in male and female rats. *Exp Clin Psychopharmacol* 16: 165–177, 2008.

Petry NM, Alessi SM, Marx J, Austin M, Tardif M. Vouchers versus prizes: contingency management treatment of substance abusers in community settings. *J Consult Clin Psychol* 73: 1005–1014, 2005.

Petry NM, Martin B. Low-cost contingency management for treating cocaine- and opioid-abusing methadone patients. *J Consult Clin Psychol* 70: 398–405, 2002.

Petry NM, Martin B, Cooney JL, Kranzler HR. Give them prizes, and they will come: contingency management for treatment of alcohol dependence. *J Consult Clin Psychol* 68: 250–257, 2000.

Petry NM, Tedford J, Austin M, Nich C, Carroll KM, Rounsaville BJ. Prize reinforcement contingency management for treating cocaine users: how low can we go, and with whom? *Addiction* 99: 349–360, 2004.

Piazza PV, Deminière JM, Le Moal M, Simon H. Factors that predict individual vulnerability to amphetamine self-administration. *Science* 245: 1511–1513, 1989.

Pierce RC, Kumaresan V. The mesolimbic dopamine system: the final common pathway for the reinforcing effect of drugs of abuse? *Neurosci Biobehav Rev* 30: 215–238, 2006.

Piray P, Keramati MM, Dezfouli A, Lucas C, Mokri A. Individual differences in nucleus accumbens dopamine receptors predict development of addiction-like behavior: a computational approach. *Neural Comput* 22: 2334–2368, 2010.

Porrino LJ, Lyons D, Smith HR, Daunais JB, Nader MA. Cocaine self-administration produces a progressive involvement of limbic, association, and sensorimotor striatal domains. *J Neurosci* 24: 3554–3562, 2004.

Porrino LJ, Smith HR, Nader MA, Beveridge TJR. The effects of cocaine: a shifting target over the course of addiction. *Prog Neuropsychopharmacol Biol Psychiatry* 31: 1593–1600, 2007.

Powell NJ, Redish AD. Complex neural codes in rat prelimbic cortex are stable across days on a spatial decision task. *Front Behav Neurosci* 8: 120, 2014.

Rabinovitch MS, Rosvold HE. A closed-field intelligence test for rats. *Can J Psychol* 5: 122–128, 1951.

Radke AK, Holtz NA, Gewirtz JC, Carroll ME. Reduced emotional signs of opiate withdrawal in rats selectively bred for low (LoS) versus high (HiS) saccharin intake. *Psychopharmacology (Berl)* 227: 117–126, 2013.

Ragozzino ME. The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Ann N Y Acad Sci* 1121: 355–375, 2007.

Ragozzino ME, Jih J, Tzavos A. Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. *Brain Res* 953: 205–214, 2002a.

Ragozzino ME, Ragozzino KE, Mizumori SJY, Kesner RP. Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behav Neurosci* 116: 105–115, 2002b.

Reddy DS. Neurosteroids: endogenous role in the human brain and therapeutic potentials. *Prog Brain Res* 186: 113–137, 2010.

Redish AD. Addiction as a computational process gone awry. *Science* 306: 1944–1947, 2004.

Redish AD. *The Mind within the Brain: How We Make Decisions and How those Decisions Go Wrong*. Oxford University Press, 2013.

Redish AD, Jensen S, Johnson A. A unified framework for addiction: vulnerabilities in the decision process. *Behav Brain Sci* 31: 415–437; discussion 437–487, 2008.

Reed SC, Evans SM, Bedi G, Rubin E, Foltin RW. The effects of oral micronized progesterone on smoked cocaine self-administration in women. *Horm Behav* 59: 227–235, 2011.

Regier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL, Goodwin FK. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA* 264: 2511–2518, 1990.

Regier PS, Carroll ME, Meisel RL. Cocaine-induced c-Fos expression in rats selectively bred for high or low saccharin intake and in rats selected for high or low impulsivity. *Behav Brain Res* 233: 271–279, 2012.

Regier PS, Claxton AB, Zlebnik NE, Carroll ME. Cocaine-, caffeine-, and stress-evoked cocaine reinstatement in high vs. low impulsive rats: treatment with allopregnanolone. *Drug Alcohol Depend* 143: 58–64, 2014.

Rescorla RA. Experimental extinction. In: *Handbook of contemporary learning theories*, edited by Mowrer RR, Klein SB. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 2001, p. 119–154.

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

Ritz MC, Lamb RJ, Goldberg SR, Kuhar MJ. Cocaine receptors on dopamine transporters are related to self-administration of cocaine. *Science* 237: 1219–1223, 1987.

Robbins TW. The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. *Psychopharmacology (Berl)* 163: 362–380, 2002.

Robbins TW, Gillan CM, Smith DG, de Wit S, Ersche KD. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci (Regul Ed)* 16: 81–91, 2012.

Robertson HA, Peterson MR, Murphy K, Robertson GS. D1-dopamine receptor agonists selectively activate striatal c-fos independent of rotational behaviour. *Brain Res* 503: 346–349, 1989.

Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev* 18: 247–291, 1993.

Robinson TE, Berridge KC. The psychology and neurobiology of addiction: an incentive-sensitization view. *Addiction* 95 Suppl 2: S91–117, 2000.

Rosenberg DR, Keshavan MS. A.E. Bennett Research Award. Toward a neurodevelopmental model of of obsessive--compulsive disorder. *Biol Psychiatry* 43: 623–640, 1998.

Ruskin DN, Marshall JF. Amphetamine- and cocaine-induced fos in the rat striatum depends on D2 dopamine receptor activation. *Synapse* 18: 233–240, 1994.

Saxena S, Brody AL, Schwartz JM, Baxter LR. Neuroimaging and frontal-subcortical circuitry in obsessive-compulsive disorder. *Br J Psychiatry Suppl* : 26–37, 1998.

Schmitzer-Torbert NC, Redish AD. Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. *Neuroscience* 153: 349–360, 2008.

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

Schmitzer-Torbert N, Redish AD. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. *J Neurophysiol* 91: 2259–2272, 2004.

Schoenbaum G, Nugent SL, Saddoris MP, Setlow B. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* 13: 885–890, 2002.

Schultz W, Tremblay L, Hollerman JR. Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology* 37: 421–429, 1998.

See RE. Neural substrates of conditioned-cued relapse to drug-seeking behavior. *Pharmacol Biochem Behav* 71: 517–529, 2002.

See RE, Elliott JC, Feltenstein MW. The role of dorsal vs ventral striatal pathways in cocaine-seeking behavior after prolonged abstinence in rats. *Psychopharmacology (Berl)* 194: 321–331, 2007.

Sheffer C, Mackillop J, McGeary J, Landes R, Carter L, Yi R, Jones B, Christensen D, Stitzer M, Jackson L, Bickel W. Delay discounting, locus of control, and cognitive impulsiveness independently predict tobacco dependence treatment outcomes in a highly dependent, lower socioeconomic group of smokers. *Am J Addict* 21: 221–232, 2012.

Shiflett MW, Brown RA, Balleine BW. Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *J Neurosci* 30: 2951–2959, 2010.

Sinha R, Fox H, Hong K-I, Sofuoglu M, Morgan PT, Bergquist KT. Sex steroid hormones, stress response, and drug craving in cocaine-dependent women: implications for relapse susceptibility. *Exp Clin Psychopharmacol* 15: 445–452, 2007.

Sinha R, Lacadie C, Skudlarski P, Fulbright RK, Rounsaville BJ, Kosten TR, Wexler BE. Neural activity associated with stress-induced cocaine craving: a functional magnetic resonance imaging study. *Psychopharmacology (Berl)* 183: 171–180, 2005.

Smith KS, Graybiel AM. A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron* 79: 361–374, 2013.

Sofuoglu M, Mitchell E, Kosten TR. Effects of progesterone treatment on cocaine responses in male and female cocaine users. *Pharmacol Biochem Behav* 78: 699–705, 2004.

Solomon RL, Corbit JD. An opponent-process theory of motivation. I. Temporal dynamics of affect. *Psychol Rev* 81: 119–145, 1974.

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

Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G. Reward prediction error signaling in posterior dorsomedial striatum is action specific. *J Neurosci* 32: 10296–10305, 2012.

Stead JDH, Clinton S, Neal C, Schneider J, Jama A, Miller S, Vazquez DM, Watson SJ, Akil H. Selective breeding for divergence in novelty-seeking traits: heritability and enrichment in spontaneous anxiety-related behaviors. *Behav Genet* 36: 697–712, 2006.

Steiner AP, Redish AD. The road not taken: neural correlates of decision making in orbitofrontal cortex. *Front Neurosci* 6: 131, 2012.

Steiner H, Gerfen CR. Cocaine-induced c-fos messenger RNA is inversely related to dynorphin expression in striatum. *J Neurosci* 13: 5066–5081, 1993.

Stewart J. Pathways to relapse: the neurobiology of drug- and stress-induced relapse to drug-taking. *J Psychiatry Neurosci* 25: 125–136, 2000.

Stitzer M, Bigelow G. Contingency management in a methadone maintenance program: availability of reinforcers. *Int J Addict* 13: 737–746, 1978.

Strain EC, Stitzer ML, Liebson IA, Bigelow GE. Comparison of buprenorphine and methadone in the treatment of opioid dependence. *Am J Psychiatry* 151: 1025–1030, 1994.

Sun H, Green TA, Theobald DEH, Birnbaum SG, Graham DL, Zeeb FD, Nestler EJ, Winstanley CA. Yohimbine increases impulsivity through activation of cAMP response element binding in the orbitofrontal cortex. *Biol Psychiatry* 67: 649–656, 2010.

Swanson LW. Cerebral hemisphere regulation of motivated behavior. *Brain Res* 886: 113–164, 2000.

Toates FM. *Motivational Systems*. CUP Archive, 1986.

Toledo-Rodriguez M, Sandi C. Stress during Adolescence Increases Novelty Seeking and Risk-Taking Behavior in Male and Female Rats. *Front Behav Neurosci* 5: 17, 2011.

Tolman EC. The determiners of behavior at a choice point. *Psychological Review* 45: 1–41, 1938.

Tolman EC. There is more than one kind of learning. *Psychological Review* 56: 144–155, 1949.

Torregrossa MM, Quinn JJ, Taylor JR. Impulsivity, compulsivity, and habit: the role of orbitofrontal cortex revisited. *Biol Psychiatry* 63: 253–255, 2008.

Trope Y, Liberman N. Temporal construal. *Psychological Review* 110: 403–421, 2003.

Vanderschuren LJMJ, Di Ciano P, Everitt BJ. Involvement of the dorsal striatum in cue-controlled cocaine seeking. *J Neurosci* 25: 8665–8670, 2005.

Volkow ND, Fowler JS. Addiction, a disease of compulsion and drive: involvement of the orbitofrontal cortex. *Cereb Cortex* 10: 318–325, 2000.

Volkow ND, Fowler JS, Wang G-J. The addicted human brain: insights from imaging studies. *J Clin Invest* 111: 1444–1451, 2003.

Volkow ND, Fowler JS, Wang GJ, Baler R, Telang F. Imaging dopamine's role in drug abuse and addiction. *Neuropharmacology* 56 Suppl 1: 3–8, 2009.

Volkow ND, Fowler JS, Wang G-J, Swanson JM, Telang F. Dopamine in drug abuse and addiction: results of imaging studies and treatment implications. *Arch Neurol* 64: 1575–1579, 2007.

Volkow ND, Wang G-J, Telang F, Fowler JS, Logan J, Childress A-R, Jayne M, Ma Y, Wong C. Cocaine cues and dopamine in dorsal striatum: mechanism of craving in cocaine addiction. *J Neurosci* 26: 6583–6588, 2006.

Walker QD, Nelson CJ, Smith D, Kuhn CM. Vaginal lavage attenuates cocaine-stimulated activity and establishes place preference in rats. *Pharmacol Biochem Behav* 73: 743–752, 2002.

Welter M, Vallone D, Samad TA, Meziane H, Usiello A, Borrelli E. Absence of dopamine D2 receptors unmasks an inhibitory control over the brain circuitries activated by cocaine. *Proc Natl Acad Sci USA* 104: 6840–6845, 2007.

Winstanley CA. The orbitofrontal cortex, impulsivity, and addiction. *Annals of the New York Academy of Sciences* 1121: 639–655, 2007.

Winstanley CA, Theobald DEH, Cardinal RN, Robbins TW. Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J Neurosci* 24: 4718–4722, 2004.

Wise RA. Neurobiology of addiction. *Curr Opin Neurobiol* 6: 243–251, 1996.

Wise RA. Dopamine, learning and motivation. *Nat Rev Neurosci* 5: 483–494, 2004.

De Wit H, Richards JB. Dual determinants of drug use in humans: reward and impulsivity. In: *Motivational factors in the etiology of drug abuse. Volume 50 of the Nebraska Symposium on Motivation*, edited by Bevins RA, Bardo MT. Lincoln, NE, US: University of Nebraska Press, 2004, p. 19–55.

De Wit H, Schmitt L, Purdy R, Hauger R. Effects of acute progesterone administration in healthy postmenopausal women and normally-cycling women. *Psychoneuroendocrinology* 26: 697–710, 2001a.

De Wit H, Schmitt L, Purdy R, Hauger R. Effects of acute progesterone administration in healthy postmenopausal women and normally-cycling women. *Psychoneuroendocrinology* 26: 697–710, 2001b.

Woolverton WL, English JA, Weed MR. Choice between cocaine and food in a discrete-trials procedure in monkeys: a unit price analysis. *Psychopharmacology (Berl)* 133: 269–274, 1997.

Yamamoto DJ, Nelson AM, Mandt BH, Larson GA, Rorabaugh JM, Ng CMC, Barcomb KM, Richards TL, Allen RM, Zahniser NR. Rats classified as low or high cocaine locomotor responders: a unique model involving striatal dopamine transporters that predicts cocaine addiction-like behaviors. *Neurosci Biobehav Rev* 37: 1738–1753, 2013.

Yin HH, Knowlton BJ. Contributions of striatal subregions to place and response learning. *Learn Mem* 11: 459–463, 2004.

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

Yin HH, Knowlton BJ, Balleine BW. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci* 22: 505–512, 2005a.

Yin HH, Knowlton BJ, Balleine BW. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav Brain Res* 166: 189–196, 2006.

Yin HH, Mulcare SP, Hilário MRF, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, Costa RM. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* 12: 333–341, 2009.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* 22: 513–523, 2005b.

Young ST, Porrino LJ, Iadarola MJ. Cocaine induces striatal c-fos-immunoreactive proteins via dopaminergic D1 receptors. *Proc Natl Acad Sci USA* 88: 1291–1295, 1991.

Zapata A, Minney VL, Shippenberg TS. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *J Neurosci* 30: 15457–15463, 2010.