Dopamine Modulation of Choice Behavior Following Unexpected Reward Omission

by

Kirsten A. Porter-Stransky

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

Doctoral Committee:

Professor Terry E. Robinson, Chair
Professor Jill B. Becker
Professor Robert T. Kennedy
Professor Martin Sarter

DEDICATION

This dissertation is dedicated to my amazing mother and lifelong best friend, Paulette C. Porter-Stransky, and to the memory of my loving grandparents, Terry and Cordula Porter. These three have continually demonstrated self-sacrificing love, support, and belief in me, which I have no doubt is the reason that I am able to accomplish my dream of earning a Ph.D.

ACKNOWLEDGMENTS

I will forever be grateful for the numerous individuals who have provided support and guidance throughout my dissertation research and grad school adventure. First and foremost, I would like to thank my primary mentor, Dr. Brandon Aragona, for choosing me to be one of his first graduate students and for providing me with so many professional opportunities. I am also very grateful for Dr. Terry Robinson and his continued mentorship and willingness to step in as my dissertation chair. Additionally, I would like to thank my other committee members, Drs. Robert Kennedy, Jill Becker, and Martin Sarter, for their ongoing support and involvement in my training. Additionally, I would like to acknowledge Drs. Peggy Gnegy, Natalie Tronson, and Kent Berridge for their mentorship. And thanks to Jillian Seiler and Daniel Castro for proof-reading this dissertation.

As the Aragona lab has grown over the past 5 years, there have been many individuals that have contributed to making the lab such a fun, productive, and intellectually stimulating environment, including Dr. Shanna Resendez, Dr. Vedran Lovic, Dr. Bryan Singer, Jillian Seiler, Cailtin Vander Weele, Curtis Austin, Ary Bryan, Piper Keyes, Natalie Nevárez, Dr. Sean Ma, and many others. I have also been blessed to have a number of fantastic undergraduate research assistants; in particular, Jillian Seiler has helped collect data for many of the experiments in this dissertation and has recently become an invaluable help as a technician and lab manager. Other great undergraduate volunteers over the years (in no particular order) include: Nathan Nartker, Ben Blackmer,

Erin Hawkins, Tom Geesey, Molly Hershman, Scott Kirsch, Diana Juncaj, Bianca Jenkins, Melanie Schweir, Jessica Grimmer, Jeff Brow, Katie Copeland, Marie Rieman-Klingler, Alber Trieu, Brianna Hatch-Vallier, Amelia Stone, Anna Nagler, and Sarah Berke, and Dan Saffaei.

A number of experiments would not have been possible if it were not for our world-class collaborators at Michigan. A few years ago, Dr. Robert Kennedy agreed to begin a collaboration with Brandon and I, and we have been collaborating on a number of projects ever since. The experience and knowledge that I have gained throughout these collaborations have significantly contributed to my development as a scientist. A special thanks to Dr. Omar Mabrouk for teaching me microdialysis and to my other colleagues and friends in the Kennedy lab including Dr. Tom Slaney, Jenny Wong, and Dr. Neil Hershey. Also, much thanks to our more recent collaborators from the Berke lab, Arif Hamid and Dr. Josh Berke, for helping us set up optogenetics in the Aragona lab.

Graduate school would definitely not have been as enjoyable if it were not for wonderful friends in the department including Shanna Resendez, Alexandra DiFeliceantonio, Lindsay Yager, Katie Goldey, Daniel Castro, Jocelyn Richard, Matt Howe, Ben Saunders, Arif Hamid, Mike Robinson, Liz Cogan, Lindsay Bowman, Anne Berry, Alaina Neary Case, Giovanna Paolone, Morgan Gustison, Howard Gritton, Jeff Petibone, Jonté Jones, Natalie Nevárez, Shelley Warlow, Shannon Cole, Alex Kawa, Katie Yoest, Kyra Phillips, Ashley Schmeling, Aneesha Badrinarayan, and the many others that I have met along the way.

Additionally, I would like to thank Marc Bradshaw for research technical support; Bakari Wooten and Buda Martonyi for keeping our lab facilities in tip top shape; Kathy Gimson and the animal care staff for providing a high quality of care to our rats; and Susan Douglas and Mary Mohrbach for administrative support. Of course, research cannot be conducted without funding: I would like to acknowledge the Rackham Graduate School for awarding me a Rackham Merit Fellowship and two small research grants, the Department of Psychology, and the Interdisciplinary Training Grant on Substance Abuse from the National Institutes of Health, Ruth L. Kirschstein National Research Service Award T32 DA007267.

Looking back in time for a moment, I would like to especially thank my undergraduate mentors from Sweet Briar College, Drs. Dan Gottlieb and Tim Loboschefski, for being incredible teachers and giving me the confidence to apply to Michigan for graduate school. My interest in research began while working in Dr. Dan Gottlieb's animal learning lab, and without his guidance and mentorship, I would not have been prepared for graduate school. Additionally, I would like to thank Dr. Cedric Williams from the University of Virginia, who allowed me to volunteer in his lab when I was an undergraduate. My experience in his lab confirmed that I wanted to become a behavioral neuroscientist.

Last, but in no way least, I would like acknowledge my amazing family. My incredible mom and dad have always been supportive and sacrificed whatever necessary for my education, especially since they were not lucky enough to have such opportunities for themselves. I will forever be indebted to my parents for their self-sacrificing love and for helping me become the person that I am. Additionally, I would like to thank my

wonderful grandparents, Cordula Porter, Terry Porter, and Gerry Stransky. Papa has always been very supportive of my education and has helped me in countless ways. My grandmother importantly taught me from a young age how to be a strong and independent woman, and, more recently, when grad school was especially challenging, Grandad still believed in me and inspired me to keep going. In fact, when diagnosed with cancer, he told his doctors that his goal was to see me get my Ph.D. Finally, I would like to thank my awesome boyfriend, Hays Watson. When we realized we could not attend the same graduate school, he decided not to let that get in our way; he has faithfully visited me in Ann Arbor every few weeks for the past five years.

Ya'll are the best, the absolute best.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1:

GENERAL INTRODUCTION

Adaptive Reward Seeking Behavior

Fundamental to the existence of animals, including humans, is the ability to obtain natural rewards, such as food and water. Securing these resources is necessary for fitness and reproduction (Pyke, 1984), and critical for successful foraging is the ability to learn which foraging patches are rich in resources. When resources have previously been found in certain patches, many animals will subsequently search in those patches again, termed an area-restricted search (Kareiva and Odell, 1987; Hills, 2006). This behavior is phylogenetically ancient, having been observed in numerous species including *C. elegans* (Hills et al., 2004), insects (White et al., 1984), rodents (Benedix, 1993; Haskell, 1997; Reid and Staddon, 1998), wandering albatrosses (Weimerskirch et al., 2007), European polecats (Lode, 2000), and humans (Hills et al., 2013; Wolfe, 2013).

Similarly, animals must be able to recognize when resource availability changes and alter motivated behavior accordingly toward a new foraging patch. Modeling a rapid depletion, if an animal is given significantly less reward than expected, they are often observed to display behavioral reactions that have been described as frustrating and aversive (Amsel, 1958; Papini and Dudley, 1997). These responses to the omission of an expected reward are also thought to be evolutionarily old, since they are observed in fish (Vindas et al., 2012), rodents (Gallup Jr, 1965; Salinas et al., 1993; Kerfoot et al., 2008),

monkeys (Tinklepaugh, 1928), and humans (Papini and Dudley, 1997; Abler et al., 2006). These activated behavioral responses may be helpful in prompting the animal to search elsewhere for necessary rewards. Indeed, being able to exhibit flexible behavior in a dynamic environment is evolutionarily adaptive.

With modern technology, humans in many societies no longer need to forage in nature for their food. However, the same cognitive and behavioral skills required for optimal foraging are utilized daily by people in visual search (Wolfe, 2013) and semantic memory (Hills et al., 2012). Indeed, goal-directed cognition likely evolved from basic foraging behaviors (Hills, 2006). Additionally, deficits in cognitive and behavioral flexibility are observed in patients with neurological conditions such as Parkinson's disease (Cools et al., 2001; Frank, 2005), Huntington's disease (Josiassen et al., 1983; Montoya et al., 2006), schizophrenia (Thoma et al., 2007; Floresco et al., 2009), and drug abuse (Aharonovich et al., 2006; Colzato et al., 2009). Therefore, investigating the behavioral and neurobiological mechanisms that mediate adaptive, reward seeking behavior under variable conditions is necessary for understanding how the brain has evolved to mediate these fundamental behaviors and paves the way for elucidating the relationship between certain disorders and the neural circuitry that underlies these behaviors.

Dopaminergic Coding of Reward

The basal ganglia, which mediate many motivated behaviors (Haber, 2003; Freeze et al., 2013), exist in some form in amniotes, amphibians, jawed fish, lamprey, and mammals, suggesting that the vertebrate, common ancestor of these species had basal ganglia (Reiner et al., 1998). Furthermore, neurons containing the neurotransmitter

dopamine (DA) have strong inputs into the striatum, and striatal DA facilitates many motivated and goal-directed behaviors (Swanson et al., 1997; Berridge and Robinson, 1998; Aragona and Wang, 2009) and exists in vertebrate and invertebrate organisms (Smeets and González, 2000; Mustard et al., 2005).

Interference with DA transmission in a number of species impairs reward seeking behaviors. For example, lesioning DA neurons impairs area-restricted search in *C. elegans* (Hills et al., 2004), and altering DA transmission can cause impairments in behavioral flexibility in rodents and humans (Frank et al., 2004; Haluk and Floresco, 2009). The nucleus accumbens (NAc) in particular receives strong dopaminergic projections from the ventral tegmental area (Ikemoto, 2007), and interfering with endogenous neurotransmission in the NAc impairs certain forms of flexible behavior (Taghzouti et al., 1985; Cardinal et al., 2001; Floresco et al., 2006a; Gill et al., 2010), although little is known about the role of DA in mediating choice behavior when an expected reward is omitted (Annett et al., 1989; Reading and Dunnett, 1991). Together, these studies suggest that DA is a likely candidate for mediating adaptive, flexible reward seeking behavior that is fundamental to numerous species.

Much of the field's knowledge about the firing patterns of DA neurons and their relationship to behavior comes from electrophysiological recordings. Recordings of putative, midbrain DA neuron have revealed that these neurons can exhibit multiple firing patterns. Specifically, they have a low, single-spike firing rate (Grace and Bunney, 1984b). Additionally, they can burst fire with multiple, consecutive spikes (Grace and Bunney, 1984a). Much attention has been given to the phasic bursts of activity, since these firing patterns correlate to reward predictive information (Schultz, 1998), whereas

much less attention is focused on the single spike, baseline firing rate of DA neurons. An iconic study by Wolfram Schultz's laboratory using monkeys has revealed that when an unexpected reward is provided, putative DA neurons burst fire to the event of receiving the reward (Schultz et al., 1997). If a conditioned stimulus predicts the reward, once subjects learn the relationship between the conditioned stimulus and the reward, DA neurons burst fire to the conditioned stimulus instead of the reward. This transition of phasic activity from the reward to the conditioned stimulus has also been demonstrated in rats (Pan et al., 2005). Finally, when an expected reward is withheld, putative DA neurons phasically decrease their firing rate (Schultz et al., 1997; Roesch et al., 2007). These phasic changes in DA neuronal activity are thought to cause changes in DA release in forebrain regions which receive strong dopaminergic projections, such as the nucleus accumbens (NAc).

As predicted by the aforementioned electrophysiology studies, experiments measuring sub-second DA transmission in the NAc have demonstrated that extra-cellular DA concentration ([DA]) phasically increases when an unexpected reward is received (Day et al., 2007; Hart et al., 2014) and decreases in response to certain negative events (McCutcheon et al., 2012). In Pavlovian conditioning, as animals gain experience with a conditioned stimulus that predicts reward, phasic increases in [DA] are observed to the cue (Day et al., 2007). These phasic increases in [DA] that occur in relation to a conditioned stimulus are termed cue-evoked DA. With overtraining, cue-evoked DA to a Pavlovian cue diminishes even though the condition approach response persists (Clark et al., 2013).

In addition to Pavlovian conditioning, cue-evoked DA in the NAc occurs during operant conditioning tasks. Under certain circumstances, the magnitude of cue-evoked DA appears to track the utility of specific options during reward seeking behavior (Phillips et al., 2007b). Specifically, once rats have learned the value of different options, greater cue-evoked DA is observed for very low effort options (but does not differentiate between two higher, but different, effort options) and for immediate versus delayed reward options (Day et al., 2010; Gan et al., 2010).

Since DA neuronal firing and [DA] in the NAc increase during unexpected reward, mesolimbic DA has been hypothesized to signal reward prediction errors to aid in reinforcement learning. In this theory, if an outcome is better than expected, a phasic increase in DA would occur, whereas if something is worse than expected, a phasic decrease in DA release would be expected to occur. If aversive stimuli are represented as "negative reward," phasic decreases in DA are predicted (Daw and Touretzky, 2002). Along these lines, phasic increases in DA transmission have been hypothesized to reinforce actions, and phasic decreases in DA transmission are thought to punish actions, making them less likely to occur in the future (Bromberg-Martin et al., 2010b). And, indeed, there is evidence that midbrain DA neuronal firing (Schultz et al., 1997), as well as phasic changes in [DA] in the NAc (Hart et al., 2014), encode reward prediction errors. However, there is also evidence that DA is not necessary for reward learning (Cannon and Palmiter, 2003; Robinson et al., 2005) and that DA transmission in the NAc is necessary for learning about rewards only when incentive salience is attributed to the reward predictive cues (Flagel et al., 2011).

Furthermore, electrophysiology studies of putative midbrain DA neurons have discovered heterogeneity in the responses of putative midbrain DA neurons to reward. While some DA neurons follow the classic model of increasing firing to an unpredicted reward and being inhibited by negative events, a large proportion of DA neurons are excited by both appetitive and aversive stimuli and outcomes (Matsumoto and Hikosaka, 2009). Similarly to the various electrophysiological reactions of midbrain DA neurons to aversive stimuli, some aversive stimuli decrease, whereas others increase, [DA] in terminal regions. Phasic decreases in [DA] occur in response to aversive taste stimuli (Wheeler et al., 2011; McCutcheon et al., 2012), and phasic increases in [DA] in the NAc have been observed during social defeat (Anstrom et al., 2009). Interestingly, both phasic increases and decreases have been observed during the presentation of cues that predict foot shock; the direction of the DA response varies by subregion of the NAc (Badrinarayan et al., 2012; Oleson et al., 2012).

The challenges of studying dopamine transmission in behaving animals

Much of our knowledge about DA signaling in relation to specific reward seeking behaviors has been obtained using electrophysiology or electrochemical techniques. Each technique can be a valuable tool to facilitate the investigation of DA's role in adaptive goal directed behavior. However, it is critical to understand the advantages as well as limitations of the techniques in order to properly determine which techniques are most suitable to address specific experimental questions and what conclusions can be made from experiments utilizing these various techniques.

Electrophysiological recordings of putative midbrain DA neurons

Much of the field's knowledge about DA's role in reward has been achieved through electrophysiology experiments. While such studies have made important contributions to the field's understanding of DA neurons, there is uncertainty with regards to which subpopulations of neurons are being studied. To unambiguously identify DA neurons and their projection targets, subjects must be anesthetized (Ungless et al., 2004; Lammel et al., 2008; Lammel et al., 2011) which is not a viable option for recording in awake, behaving animals. Therefore, in awake animals, putative DA neurons have traditionally been identified by a waveform pioneered by Grace and Bunny (Grace and Bunney, 1984a, b). Although frequently used to identify putative DA neurons, such criterion has led to misidentification of neurons (Ungless et al., 2004; Brischoux et al., 2009). Furthermore, activity patterns of DA neurons may change depending on subject's age (Margolis et al., 2006a; Grace et al., 2007).

Importantly, even if such electrophysiological criteria do accurately identify DA neurons, these criteria may only identify a certain subpopulation of DA neurons. For example, in one study only about 14% of the 258 neurons recorded from in the ventral tegmental area (VTA) and substantia nigra (SN) met this established eletrophysiological criteria for DA neurons (Roesch et al., 2007), and in another experiment only about 22% (Nishino et al., 1987), although anatomical studies indicate that approximately 55% of VTA neurons and at least 88% of SN neurons are dopaminergic (Margolis et al., 2006a).

Similarly, many experimenters have identified putative DA neurons based on a large hyperpolariztion-activated cation current (I_h), and cells lacking this current have been classified as nondopaminergic (Bonci and Malenka, 1999; Margolis et al., 2003;

Margolis et al., 2006b; Wanat et al., 2008), although the presence of I_h has not always reliably predicted TH labeling (Cameron et al., 1997; Jones et al., 2000; Margolis et al., 2003; Margolis et al., 2006a). Recent work using retrograde tracing methods has demonstrated that there are subpopulations of DA neurons that have been neglected. Indeed, there are differences in the characteristics of DA neurons including the magnitude of I_h which can significantly vary among DA neurons with different projection targets (Lammel et al., 2011). DA neurons located in the ventromedial posterior VTA projecting to the medial NAc in particular have been ignored since these neurons do not fit the traditional electrophysiological characteristics of DA neurons, since they have very small I_h and inward leak currents (Lammel et al., 2011).

Indeed, DA neurons are not homogenous; rather, there are distinct subpopulations that project to different forebrain terminal regions (Ikemoto, 2007). Relying on spike parameters and I_h to identify DA neurons has likely biased interpretations of how DA neurons respond in behavioral contexts (Garris and Rebec, 2002; Borgkvist et al., 2011). For example, while it is well established that midbrain DA neurons respond to rewards (Schultz et al., 1997; Hyland et al., 2002b; Pan et al., 2005; Roesch et al., 2007), recent evidence demonstrates that DA release to unexpected rewards is not uniform through the striatum and that DA is differentially released in specific subregions of the dorsal and ventral striatum (Brown et al., 2011). This is contrary to the notion that DA globally increases throughout the striatum to unexpected rewards and reward predictive cues (Schultz, 1998). As most electrophysiology studies primarily capture classical DA neurons, more unconventional DA neurons remain unstudied. In fact, very little is known about these unconventional DA neurons projecting to the medial shell. Understanding

the role of DA in these subregions of the ventral striatum is important, since the medial NAc shell can have a different role than the NAc core in certain forms of behavioral flexibility (Floresco et al., 2006a). Additionally, drugs of abuse can target this subpopulation of DA neurons. For example, cocaine preferentially increases DA transmission in the medial NAc shell (Aragona et al., 2008; Aragona et al., 2009), and cocaine administration increases AMPAR/NMDAR ratio in these VTA neurons (Lammel et al., 2011). Therefore, employing techniques that can parse apart these subpopulations of DA neurons and their different projection targets is critical for being able to study how these different subpopulations signal specific components of reward, aversion, and motivated behavior.

Since these subpopulations of DA neurons cannot yet be reliably identified based upon their electrophysiological properties, currently, the only way to study these projection systems with certainty in awake, behaving animals is by recording in the forebrain terminal regions. Even if new innovations allow for unambiguous identification of DA neurons and their projection targets based on their electrophysiological properties, recording action potentials from the cell bodies does not speak to DA release at the cells' projection targets. While electrophysiological recordings of putative midbrain DA neurons can quantify number and frequency of action potentials, they cannot unequivocally predict DA release or account for terminal modulation. Indeed, central acetylcholine and opioid systems modulate DA release locally and independently of impulse traffic (Cragg, 2006; Britt and McGehee, 2008; Cachope et al., 2012), and drugs of abuse can alter quantile size and the probability of exocytosis of DA containing vesicles (Sulzer and Pothos, 2000; Sulzer, 2011). Therefore,

to capture the dynamics of DA transmission, phasic changes in extracellular [DA] must be measured in the terminal regions.

Measuring DA transmission in terminal regions

An alternative to recording the firing properties of cells using electrophysiology is to utilize electrochemical techniques to measure [DA] in the terminal regions to which DA neurons project. Since DA transmission can signal through volume transmission (Arbuthnott and Wickens, 2007; Agnati et al., 2010), released DA escapes the synapse and can be measured by probes in the extracellular space (Robinson et al., 2003; Kennedy, 2013).

Microdialysis

Microdialysis has been an important tool used for decades to measure changes in [DA] in distinct forebrain regions. The technique works by pumping artificial cerebrospinal fluid (aCSF) into a probe that contains a semi-permeable membrane in the brain. Neurotransmitters and other compounds near the probe diffuse into the membrane, and the resulting fractions are collected by the experimenter (Fig. 1.1A). Then, compounds within the dialysate are separated using high performance liquid chromatography (HPLC) and quantified, traditionally, by electrochemical detection since DA is easily oxidized. Microdialysis is a useful tool for measuring changes in neurotransmission over long periods of time (minutes to hours) in awake, behaving animals; however, technical limitations have obscured the precise nature of phasic DA transmission *in vivo* during motivated behavior. Since microdialysis measures compounds through analyzing dialysate extracted from the brain, and a certain volume of dialysate is required for analysis, the technique has generally been limited to sampling

changes in [DA] every 5-20 min. Therefore, traditional microdialysis can be used to study changes in tonic DA levels, but it does not have the temporal resolution to align changes in [DA] to brief behavioral or environmental events. This severely limits the experimental questions that can be addressed. For example, to use microdialysis to examine [DA] in relation to cocaine cues and drug-seeking behavior, experimental designs have been altered in attempt to differentiate these events (Ito et al., 2000). However, it has now been established that there are temporally and functionally distinct phasic DA signals in relation to different components of drug self-administration that cannot be captured with microdialysis (Stuber et al., 2005).

Moreover, microdialysis captures [DA] in relatively large areas since the probes generally span millimeters in length and hundreds of microns in diameter. As such, the resulting fractions result in one data point per several minutes, and the measure spans a large area of terminal inputs. In the dorsal striatum it is known that the arborization of just one DA neuron is quite extensive covering over 1 mm^3 (Arbuthnott and Wickens, 2007). Under these conditions, using a large probe is understandable. However, the NAc core, NAc shell, and olfactory tubercle in the ventral striatum receive inputs from different populations of DA neurons (Ikemoto, 2007). Indeed, evidence is accumulating that there are distinct subpopulations of DA neurons with differing electrophysiological characteristics and projection targets which differentially react to rewards, aversive stimuli, and drugs of abuse (Ikemoto, 2007; Lammel et al., 2008; Matsumoto and Hikosaka, 2009; Bromberg-Martin et al., 2010b; Brown et al., 2011; Roeper, 2013; Lammel et al., 2014). Since subregions of the ventral striatum are more discrete than of

the dorsal striatum, smaller probes are needed to reliably measure changes in [DA] downstream from the different DA neuron subpopulations.

Recent technological advances have greatly improved the temporal and spatial resolution of microdialysis (Schultz and Kennedy, 2008). The length of sampling probes can be reduced to one millimeter or less (Kennedy, 2013). In fact, using push-pull perfusion, probes can measure specific spatial subregions as small as 0.004 mm^3 (Slaney et al., 2013). A potential issue in microdialysis is finding the appropriate compromise between spatial and temporal resolutions for the experimental question. While small dialysis probes can measure from a specific subregion of the brain, flow rates must be lower which requires longer sampling periods for each fraction (20 min). The temporal resolution can be greatly improved with larger probes, but the inherent tradeoff is that one is sampling from a larger area in the brain. Additionally, when investigating neurotransmission during reward seeking behavior, it is important that sampling equipment not interfere with the animal's behavior. Therefore, output lines must be long enough to extend outside the operant chamber. Fractions within these lines can begin to blend together because of diffusion, flow, and different concentration gradients over time. A solution to this problem is segmenting flow, whereby droplets of oil separate adjacent plugs; using this technique, temporal resolution increase to the order of seconds (Wang et al., 2008; Wang et al., 2011). Importantly, this technique can be utilized in awake, freely moving animals (Wang et al., 2011). However, subjects are usually tested in Raturns to keep the multiple fluid lines untangled; so while this setup works well for certain experimental designs, it is not yet suited for operant behavior.

Utilizing mass spectrometry (MS) instead of electrochemical detection to quantify the amounts of neurotransmitter within samples has greatly improved the levels of detection and requires less volume of dialysate per fraction (Zhang et al., 2007; Nandi and Lunte, 2009; Song et al., 2012). The separation requirements of MS are more lenient than in electrochemical detection (Kennedy, 2013). A particularly useful method utilizing MS for measuring DA in behaving animals with relatively small probes (1mm) and 60 sec temporal resolution has been developed by Robert Kennedy's lab. In this technique, small molecule neurotransmitters (such as DA) are derivatized with benzoyl chloride (Fig. 1.1A), which increases sensitivity and makes them more hydrophobic so they can better be separated during the chromatography (Song et al., 2012). This method is revolutionary in that it not only has good temporal and spatial resolution, but it allows the analysis of nearly 20 neurotransmitters and metabolites from each dialysate fraction (Fig. 1.1B). Furthermore, each fraction takes only 6-8 min to analyze, so this technique can accommodate the large number of samples necessary in behavioral studies. Additional benefits of this technique are that stable-isotope labeled internal standards can be included in each fraction to aid in quantification, and minute-by-minute changes in neurotransmission can be quantified over long periods of time.

Recent technological advances have greatly improved the spatial and temporal resolution of microdialysis. However, phasic DA transmission and behaviorally relevant events occur on the sub-second timescale; therefore, microdialysis is not yet equipped for this temporal precision. Understanding how phasic DA neurotransmission codes salient environmental events in behaving animals requires a technique that can measure changes in [DA] in specific striatal subregions on a sub-second timeframe.

Fast-scan cyclic voltammetry

The electrochemical technique fast-scan cyclic voltammetry (FSCV) at carbon fiber microelectrodes can capture sub-second changes in extra-cellular [DA] in specific terminal regions, and therefore is a valuable tool for examining phasic changes in DA transmission during motivated behaviors. In FSCV, DA molecules adhere to the carbon-fiber surface of the microelectrode because the electrode is held at a negative holding potential (-0.4 V) between measurement scans (Heien et al., 2003). A triangular waveform ramping from -0.4 V to +1.3 V at 300 V/s is applied vs. a Ag/AgCl reference electrode once every 100 ms [Fig. 1.2A; modified from (Vander Weele et al., submitted)]. The application of this measurement scan generates changes in electroactive compounds (in addition to DA) near the carbon surface. Since measurements are made ten times per second, an enormous amount of data is collected. To properly assess the data, it is helpful to view multiple measurements during one visual inspection. To facilitate this, color plots are generated in which changes in current during the scans are plotted in false color against the applied potential during over time (Fig. 1.2B). Changes in current due to the oxidation and reduction of a specific chemical compound, such as DA, can be identified based on the shape of the current by voltage plots, termed cyclic voltammograms (CVs; Fig. 1.2C), since various compounds have unique background-subtracted cyclic voltammograms (Heien et al., 2003). DA oxidizes at approximately +0.65V during the oxidative scan, and DA ortho-quinone reduces back to DA at approximately -0.2 V during the reductive scan, which creates DA's characteristic CV (Fig. 1.2C-ii). Changes in pH also affect the recorded current measured by FSCV and have a distinct CV (Fig. 1.2C-iii).

FSCV is inherently a differential technique. Unlike microdialysis, which extracts neurotransmitters from the brain to quantify them, FSCV oxidizes and reduces compounds still in the brain, revealing changes in concentrations, which is why current traces are presented as Δ[DA]. The changes in current from voltage ramps are stable over time and therefore can be subtracted out (Fig. 1.2D) to reveal the changes in current resulting from catecholamines, pH, and other compounds that can be oxidized within these parameters. Specific changes in current due to one neurotransmitter, such as DA, can be extracted and converted into changes in concentration using chemometrics (Heien et al., 2004; Keithley et al., 2010).

While other electrochemical techniques, such as amperometry, have superior temporal resolution to FSCV, there is poor chemical selectivity with amperometry since a constant potential is applied and therefore any analyte that oxidizes at the applied potential changes the recorded current (Dugast et al., 1994). In addition to positive identification of DA and excellent temporal resolution, primary advantages of FSCV include good spatial resolution and sensitivity to phasic increases and decreases in DA. Since the recording carbon fiber electrodes are approximately only 7 μm in diameter, significantly less tissue damage results from carbon fiber microelectrodes compared to microdialysis probes; in fact, tissue damage at the recording site of carbon fiber microelectordes is undetectable with light microscopy (Khan and Michael, 2003; Peters et al., 2004). Importantly, the small size of the recording electrode facilitates targeting precise striatal subregions. Therefore, rapid DA transmission in the various projection targets of midbrain DA neurons, which have been shown to be functionally distinct, can be studied.

While microdialysis examines changes in extracellular [DA] over minutes, the temporal resolution of FSCV allows DA release and reuptake to be resolved. This is clearly demonstrated using electrical stimulation of midbrain DA neurons. The abrupt increase in [DA] due to electrical stimulation is release dominated and, therefore, considered a measure of DA release. While the change in [DA] following stimulation of DA neurons is release dominated, DA reuptake via the DAT is also simultaneously occurring (Garris and Rebec, 2002)]. Indeed, when rats are pretreated with the DA uptake inhibitor nomifensine to slow DA reuptake, a four pulse stimulation delivered to the medial forebrain bundle yields an increase in [DA] in the NAc core that is four times as much as the increase in [DA] elicited by one pulse (Garris et al., 1994), demonstrating that FSCV measurements can detect DA release events of magnitudes proportional to the size of experimentally delivered stimulation of DA neurons.

Importantly, through use of a commutator, FSCV can be performed in rats during operant conditioning (Phillips et al., 2003; Day et al., 2007; Day et al., 2010). Based upon its good temporal and spatial resolution, FSCV is an excellent technique to examine phasic changes in extracellular [DA] in relation to behaviorally relevant events.

Role of dopamine in modulating choice behavior following unexpected reward omission

Many of the traditional electrophysiology studies examining putative DA neuronal firing in relationship to behaviorally relevant stimuli and outcomes have utilized simple Pavlovian tasks or tasks in which only one response is available (Niv and Schoenbaum, 2008). While Pavlovian associations and simple responses to stimuli can be very important for learning, many behaviors that human and non-human animals

perform each day involve choices and decisions among the available options. Therefore, we sought to use an instrumental task that provided subjects with the choice between two options.

As stated earlier, animals must employ adaptive foraging strategies to acquire food, but how do animals make decisions about where to forage and when to explore other foraging options? There is still debate as to how individuals choose specific actions and when to try new strategies, but a number of theories have been proposed. In basic decision making situations such as this, the field of reinforcement learning has suggested that the possible actions first are evaluated, then an action is chosen, and, finally, the chosen action can be reassessed based upon the outcome (Daw and Doya, 2006). This would be an example of model-based learning, since a goal-directed decision is made based upon one's knowledge of the environment. In contrast, model-free models track the success of past outcomes to generate rules for future actions (Dayan and Berridge, 2014). Choosing the most valuable option the majority of the time is often adaptive; however, this could result in missing out on a previously unknown, more advantageous option (Baudonnat et al., 2013). Indeed, thriving in a dynamic environment requires periodic reassessment of the available options and outcome contingencies.

To study this in the laboratory, we adapte an operant choice task in which rats can "forage" from two "patches" (i.e. two spatially distinct levers). Once trained on the task, the response outcome contingencies can be altered to examine how subjects alter their behavioral strategy to obtain reward. In chapter 2, we show that when the reward following a correct response on one lever is omitted, subjects quickly develop a choice preference for the rewarded lever. These results replicate using both male and female

rats and across all stage of the estrus cycle, demonstrating that this is a valid paradigm for examining choice preference due to manipulations of reward availability (Porter-Stransky et al., 2013).

After establishing a robust behavioral model for examining changes in behavior due to the omission of an expected reward, we sought to elucidate how DA transmission in the NAc signals this salient event. A computational modeling study has proposed that a phasic decrease in [DA] occurs to the omission of an expected reward but that there are subsequent increases in tonic levels of [DA] following reward omission (Daw and Touretzky, 2002). Additionally, DA transmission measured over hours by microdialysis has revealed increased [DA] when environmental contingencies change and subjects need to update their strategy. Therefore, after establishing the behavioral model, we tested these hypotheses by examining dopamine transmission dynamics in the NAc on two different time scales during the reward omission task. To examine change in tonic levels of extracellular DA in the NAc during reward omission, we utilize rapid-sampling microdialysis in which fractions are collected every 60 sec and analyzed with HPLC-MS (Song et al., 2012; Vander Weele et al., submitted). Then, to examine the sub-second, phasic components of DA transmission, we utilize FSCV at carbon fiber microelectrodes in the NAc during the same behavioral task.

Finally, we sought to determine the DA receptor subtype within the NAc that mediates the behavioral preference for the rewarded option during the reward omission test. DA receptors are divided into two families, the D1-like family and the D2-like family. Modeling data has shown that phasic decreases in [DA] should cause decreased binding of DA at both D1-like and D2-like DA receptors (Dreyer et al., 2010). However,

since D2-like receptors have a higher affinity for DA (Richfield et al., 1989; Marcellino et al., 2012), phasic decreases in [DA] are hypothesized to preferentially affect D2-like receptors (Bromberg-Martin et al., 2010b; Dreyer et al., 2010). To test this hypothesis, we site-specifically infused D1- and D2-like DA receptor agonists and antagonists into the NAc core prior to the first session of reward omission. To ensure that locomotor effects of the drugs did not alter the results, we conducted a locomotor experiment in which drug-naïve rats received the highest dose of each D1- and D2-like agonist and antagonist site-specifically into the NAc core.

This series of experiments establishes an ecologically valid operant paradigm for studying foraging behavior under controlled conditions in the laboratory. Using this paradigm, we show differential phasic and tonic DA transmission dynamics in the NAc during reward omission. Furthermore, we demonstrate that the behavioral preference for the rewarded option is mediated by a reduction in D2-like receptor tone in the NAc.

Figures

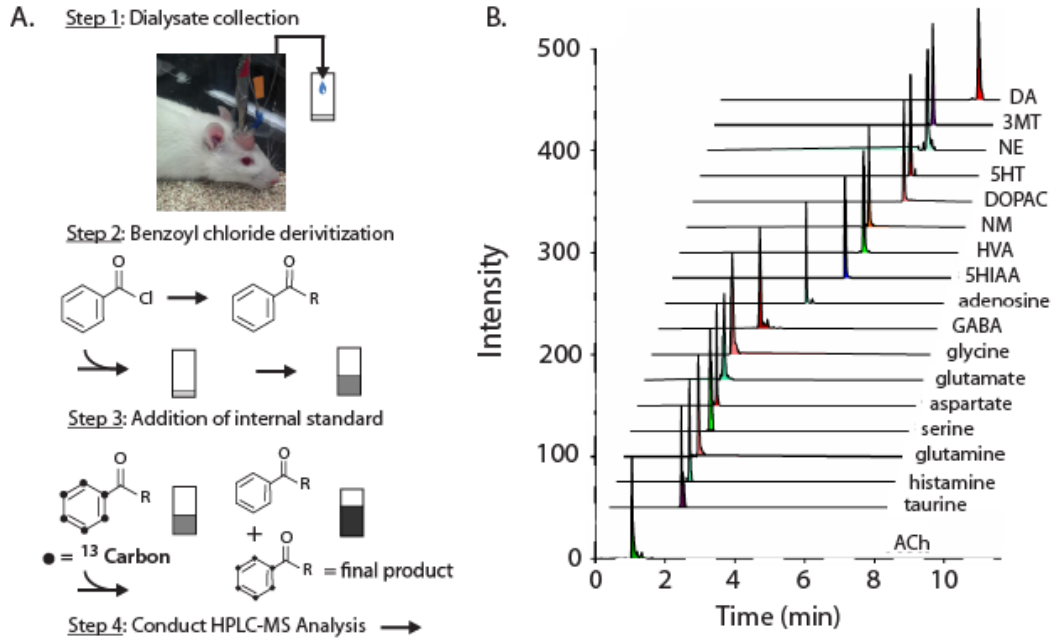

Figure 1.1. Rapid *in vivo* microdialysis sampling coupled with high perforamnce liquid chromatography and mass spectrometry detection (HPLC-MS). **A)** Dialysate was collected every min at a rate of 2.0 µL/min. Dialysate was then reacted with benzoyl chloride, and internal standards were added to the mixture for improved quantification. **B)** A representative total ion chromatogram of a dialysate sample reveals that a tremendous numbers of analytes can be measured using this technology compared to traditional microdialysis. (Figure modified from Vander Weele, Porter-Stransky, et al. submitted.)

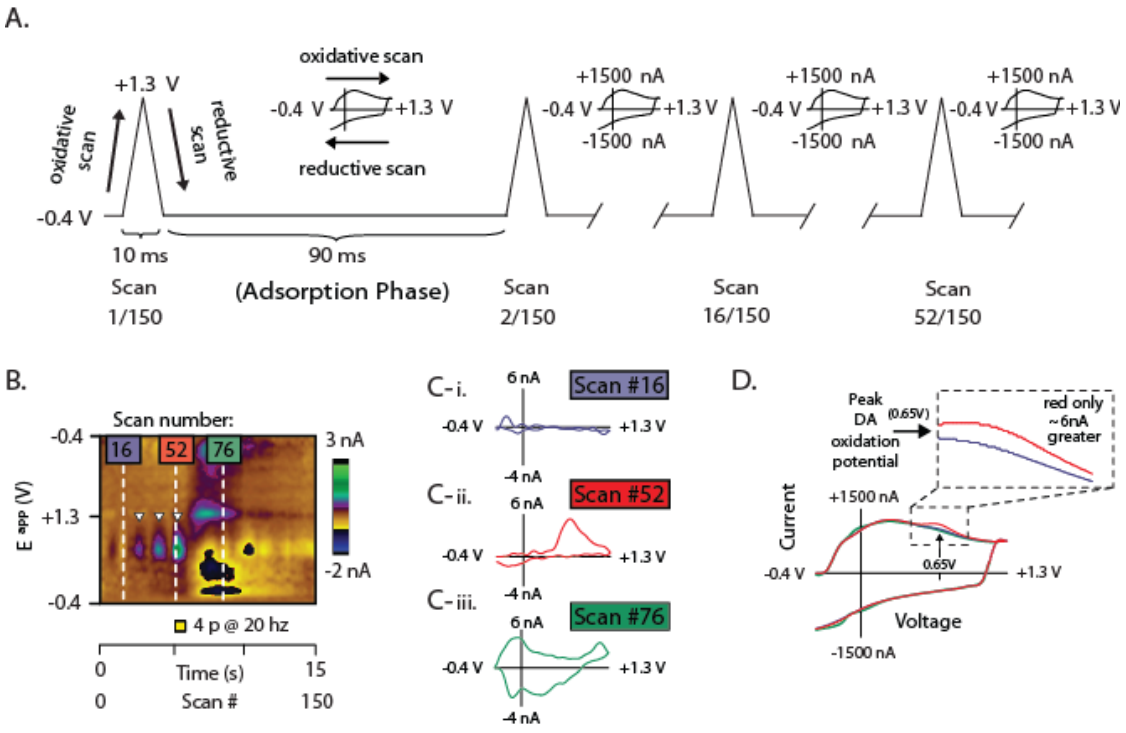

Figure 1.2. Sub-second DA transmission measured with FSCV. **A)** The carbon-fiber microelectrode is held at -0.4V for 90 ms between voltage ramps. This is referred to at the adsorption phase, because the negative potential attracts positive electroactive analytes, such as DA, and causes them to adsorb to the carbon surface. At a rate of 10 Hz, the holding potential is rapidly ramped (400 V/s) to a positive voltage (1.3V; which we refer to this as the 'oxidative scan') and then back down to the negative holding potential (-0.4V; referred to as the 'reductive scan'). This triangular scan takes ~10ms and produces a robust increase in current at the carbon surface, referred to as the charging current. Each scan, which has a corresponding CV, (in C) is represented along the x-axis (150 measures – 150 CVs - in 15 s). Highlighted are several representative scans and the resulting current changes which they cause. The charging current is extremely stable which allows for FSCV data to be background subtracted. This process permits for measures in acute current surges beyond that of the charging current to be clearly detected. **B)** Background subtracted changes in current measured during the triangular ramp are plotted in false color across the change in voltage associated with the ramp (-0.4V to 1.3V back to -0.4V; plotted on a straight line along the y-axis). C**)** Unique CVs from three different ramps (from scans 16, 52, and 76) are shown (C - i to - iii). The blue scan (C – i) shows a typical charging current, after background subtraction, (taken from a location in the color plot where there was not an obvious change in color). Conversely, the red scan (C – ii) shows the current at scan 52 which is 200 ms after the DA neurons in

the VTA received electrical stimulation (yellow box). The CV reveals an increase in current caused by an increase in DA concentration at the recording site. Electrical stimulation also elicited an expected, delayed decrease in current attributed to a basic shift in pH. The CV for this is shown by scan 76 in C – iii. **D)** Each voltage ramp generates a robust charging current; therefore, because of its consistency, it can be subtracted out to reveal the small changes in current due to oxidation and reduction of DA and other electroactive compounds. (Figure modified from Vander Weele, Porter-Stransky, et al. submitted.)

CHAPTER 2:

LABORATORY MODEL OF FORAGING BEHAVIOR, INCLUDING MANIPULATIONS OF REWARD AVAILABILITY

Introduction

Motivated behavior, such as foraging, is necessary for survival and reproductive goals (Kelley and Berridge, 2002; Aragona and Wang, 2009; Becker, 2009) and is paramount for fitness (Pyke, 1984; Stephens and Krebs, 1986). Since food availability in nature is highly dynamic, flexibility in reward-seeking behavior is critical for survival. For example, when resource availability depletes, animals must be able to recognize this alteration and rapidly adjust their behavior accordingly.

Dopamine (DA) has been shown to be involved in motivated, goal-directed behaviors (Berridge and Robinson, 1998; Wise, 2004; Salamone and Correa, 2012) and has also been proposed to be important in modulating behavioral flexibility (Haluk and Floresco, 2009; Beeler et al., 2014). However, to examine the role of DA in altering behavior when reward availability unexpectedly changes, we first had to establish a behavioral model whereby subjects had choices when "foraging" for food. We modified an instrumental task in which subjects could choose from two different levers (Day et al., 2010; Gan et al., 2010; Day et al., 2011; Sugam et al., 2012; Sugam et al., 2013).

Here, we examined two negative changes in reward availability. Initially, both levers would be equally reinforced. Then, to model changes in reward availability

(Papini and Dudley, 1997), we altered the magnitude of reward available on one lever, while the other lever remained reinforced as it had been throughout training. In one group, the reward available on one lever was reduced from 2 reward pellets to 1 (a reduction of 50%), while in the other group, reward was reduced from 1 pellet to 0 (a reduction of 100%). While both manipulations were a reduction in 1 reward pellet, subjects in the latter group developed a quicker and stronger behavioral preference for the optimal choice.

Since sex differences have been reported in a number of rodent behavioral tasks (Van Haaren et al., 1990; Jonasson, 2005; Becker and Taylor, 2008; Dalla and Shors, 2009; Sutcliffe, 2011), and should be examined in new behavioral models (Becker et al., 2005; Beery and Zucker, 2011), we used both male and female rats and monitored estrous cycles. However, we did not find any statistically significant sex differences or effects of estrous cycle in this task.

Materials & Methods

Subjects

A total of 36 Sprague-Dawley rats between 57-64 days of age (males 251-275 g and females 176-200 g) were used in these experiments. Rats were obtained from Charles River Laboratories (Winington, MA, USA) were pair-housed with a same-sex cage-mate in transparent plastic cages with metal tops. Animals were kept on a 12:12 hr reverse light-dark cycle. Experiments were run daily between 9:00 and 17:00 during the dark phase.

Mild food restriction was employed to train rats to lever-press for the food reward. Since rats naturally continue growing, daily feeding accounted for natural growth over time, which was important to maintain consistent motivation levels throughout the experiment. Subjects were food restricted to ~90% of their free feeding weight accounting for natural growth (Baker et al., 2012). Natural growth curves for free-feeding male and female rats were obtained from Charles River Laboratories (Winington, MA). After the operant session each day, rats were weighed and fed based on their weight between 15:30 – 16:30 each day during the dark cycle. Rats had free access to water in their home cages. Subjects experiencing reward reduction (described below) were fed less than subjects experiencing reward omission to equate the motivational states of the two groups, since reward reduction rats earned larger rewards than reward omission subjects.

Behavioral Paradigm

Behavioral training was conducted in MED-Associates chambers (Georgia, VT) modified locally by Marc Bradshaw at the University of Michigan. Each chamber was equipped with two cue lights, a pellet dispenser, a reward port, a white noise generator, and two Coulbourn (Whitehall, PA) levers. The reward port was centrally located, equidistant between the two levers (see Fig. 2.1A). The food reward used throughout the experiments was 45 mg BioServ chocolate-flavored dustless precision reward pellets (Bio Serv, Frenchtown, NJ).

Initially, subjects received two magazine training sessions, in which 25 reward pellets were delivered throughout the session with the inter-trial-interval (ITI) varying 40-80 sec. Then, rats learned to press two spatially distinct levers (see Fig. 2.1) to earn

up to 50 reward pellets on each lever in 1 hr. Once subjects earned 100 reward pellets (50 on each lever) in less than 60 min for two consecutive sessions (mean number of sessions = 5.2 ± 0.7), the next phase of training began in which subjects learned to discriminate between the cue lights. During these trials, one of the two cue lights would illuminate, and 5 sec later both levers would extend into the chamber for 15 sec or until 1 lever was pressed. If the lever under the illuminated cue light was chosen, a reward was delivered into the food receptacle 2 sec later. Choosing the non-illuminated lever was defined as an error; the cue light would turn off, levers would be retracted, and no food pellet would be delivered. These sessions contained 100 trials (50 with each cue light). Once subjects completed two consecutive sessions with at least 90% accuracy (mean number of sessions = 5.7 ± 0.6), they progressed to the final behavioral task described below.

Consistent with previous studies (Day et al., 2010; Gan et al., 2010; Day et al., 2011; Sugam et al., 2012), the operant paradigm contained two trial types, termed "free choice" and "forced choice" trials[1] (Fig. 2.1A&C). A cue light above the levers signaled which lever, if chosen, would yield reward. Five sec after one or both of the cue light illuminated, both levers extended into the behavioral chamber. During free choice trials, both cue lights illuminated and a response on either lever yielded reward (Fig. 2.1A), whereas on forced choice trials, subjects would receive the reward only if they chose the lever below the illuminated cue light (Fig. 2.1C). Pressing the non-illuminated lever counted as an error and no reward was delivered.

[1] In this chapter and in chapter 5, we use the term "forced choice" to describe these trials, since this is the terminology used in the literature (Day, et al., 2010; Day, et al., 2011; Sugam, et al., 2012; Sugam, et al., 2013) and how this data was originally published (Porter-Stransky, et al., 2013). However, we now recognize that this terminology can be confusing so in chapters 3-4, we call these trials "cued choice" trials instead.

One third of each session's trials (30 trials) were free choice trials, and an equal number of forced choice trials for the left and right levers (30 of each) were given in each session. Trial types were interspersed throughout each session, and the ITI varied from 10 to 30 sec. Each subject received a 1 hr training session per day containing 90 trials. Consistent with previous work (Day et al., 2010), the schedule of reinforcement progressed from FR1 to FR2 to FR4 across sessions. To advance to each FR schedule, subjects must have completed all trial types with at least 90% accuracy for two consecutive sessions. Average number of sessions to progress from FR1 to FR2 was 3.1 ± 0.4 sessions, and mean number of session to advance from FR2 to FR4 was 2.5 ± 0.3 sessions.

Throughout training, some subjects (whom later experienced reward omission) always received one reward pellet following a correct response. Others (whom later experienced reward reduction) always received two pellets after a correct operant response. Rewards were given in this way so that both conditions would be a reduction of 1 reward pellet. The fact that some subjects received more food reward during the operant sessions was accounted for in daily feeding to maintain equivalent motivational states among all subjects (see subjects section above). Regardless of whether subjects earned one or two reward pellets per trial, all subjects readily consumed their rewards throughout the session, showing no evidence of satiety.

Once stable responding, defined as a minimum of 3 consecutive days with at least 90% accuracy on each trial type, occurred on the FR4 schedule of reinforcement, subjects experienced one of two negative contingency switches. During the negative contingency switch sessions, the reward normally resulting from a correct operant response on one

lever was either reduced by 50% or completely omitted. Conversely, the other lever remained reinforced on the same schedule (i.e. its contingency was unchanged). In both reduction and omission manipulations, reward was reduced by 1 reward pellet. In other words, for subjects that received 2 reward pellets for a correct operant response during training, a response on one lever was reduced to 1 reward pellet during reward reduction (n = 5 males and 12 females). For subjects that received 1 reward pellet for each correct operant response during training, during reward omission a response on one lever yielded 0 reward pellets (n = 6 males and 13 females). More female rats were tested than male rats to ensure we tested females across all stages of the estrous cycle (see supplemental methods for the monitoring of estrous cycle).

Whether the lever that ceased to be reinforced was the right or left lever was counterbalanced across subjects, and this had no consequence on the results (data not shown). While no statistically significant lever bias was observed, if an individual rat tended to have a lever bias, the "biased lever" was chosen to be the one in which responding led to altered response contingencies (i.e. reward reduction or omission). This ensured that any changes in behavioral preference were due to the contingency switch manipulation and not a potential underlying individual lever bias. Behavior during the contingency switches did not differ between rats with no prior lever bias and those with a trending bias before the switch (data not shown).

After the 3 sessions of reward reduction or reward omission, male subjects received 6 post switch sessions. These sessions contained free choice trials and forced choice trials, and both levers were once again equally reinforced, identically to the sessions prior to the contingency switches. These extra sessions allowed us to test the

longer-term effects of reward reduction and omission on behavioral preference. Specifically, if the behavioral preference for the optimal choice was due simply to learning which lever yielded greater reward, when both levers were once again equally reinforced, subjects would be expected to choose them equally (just like they did prior to the contingency switch). However, if the lever yielding reduced or no reward was tagged with lasting aversive properties, the worse-choice lever would be expected to be avoided even when both levers were equally reinforced again.

Analysis of Behavior

Performance on the free choice and forced choice trials was automated by Med Associates software (Georgia, VT). To compare choice behavior between the reward omission and the reward reduction groups, preference scores were calculated for each rat (percentage of free choice trials choosing the optimal choice minus the percentage of free choice trials choosing the worse choice, i.e. the smaller reward lever in reward reduction and the omitted reward lever in reward omission).

Behavior during a baseline session of training, as well as the first sessions of reward reduction and reward omission, were recorded onto DVDs. Videos were scored using Behavior Tracker software to determine the locations of subjects throughout the sessions. For scoring purposes, the behavioral chambers were divided into four equal-size quadrants plus another portion of the chamber above the reward port into which rats sometimes climbed (see Fig. 2.1A for partitioning of quadrants). One quadrant was directly in front of each lever and corresponding cue lights. The percentage of time rats spent in each quadrant was analyzed.

Individuals scoring the videos were blind to the experimental manipulations and hypotheses. To maintain a high level of inter-rater reliability, certain videos were scored by multiple raters. Behavioral scoring was very consistent (above 96%), with any measure varying no more than 4% across raters.

Monitoring of estrus cycle

Vaginal lavage was performed routinely on female rats during the dark phase to track estrous cycle stage (Becker et al., 2005). An eyedropper gently inserted into the vaginal canal injected and withdrew approximately 0.3 ml saline. The eyedropper was rinsed with distilled water before each lavage to prevent cross-contamination among samples. The vaginal cells in the saline solution were viewed under a light microscope and documented by standards previously described (Tropp and Markus, 2001; Becker et al., 2005). Specifically, smears containing predominately cornified cells were classified as estrus. Smears containing a mixture of cornified and leukocytes were classified as metestrus. Smears containing primarily leukocytes were classified as diestrus, and smears containing predominantly nucleated epithelial cells were classified as proestrus. Consistent with previous experiments (Tropp and Markus, 2001), the relatively mild level of food restriction did not prevent the female rats from cycling. Female and male rats were trained identically and given the negative contingency switch once they exhibited stable behavioral responding on the operant task.

Statistics

Statistical analyses were done using SPSS Statistics 19 (IBM, Armonk, NY), and data were graphed using GraphPad Prism version 5.0 (San Diego, CA). Statistical significance for all statistical tests was defined with an α level of 0.05. Bonferroni

corrections were applied to post-hoc tests to reduce the risk of Type I errors (Sarter and Fritschy, 2008).

Consistent with previous behavioral and pharmacological studies (Haluk and Floresco, 2009; Day et al., 2010), two-way (multivariate) analyses of variance (ANOVAs) were used to examine behavioral data during baseline sessions, the contingency switch sessions, and post switch sessions as well as to screen for sex differences. Metestrus and diestrus data showed no statistically significant differences and therefore were combined for analysis (Lynch et al., 2000). The estrous cycle stage was included as a covariate in analyses (Girard and Garland Jr, 2002; Pawluski et al., 2006) to determine if it modulated the development of behavioral preferences.

Results

Establishing behavioral preference for the optimal choice during reward reduction and omission

Rats were initially trained to press two levers that yielded equal reward. During one-third of the trials termed "free choice trials," cue lights above both levers illuminated and subjects could earn a reward pellet by pressing either lever (Fig. 2.1A). During these trials, subjects earned rewards from both levers, and showed no reliable preference for one lever over the other during free choice trials (Fig. 2.1B; $t_{(10)} = 1.489$, $p = 0.167$), which was expected since both levers were equally rewarded. Conversely, during two-thirds of the trials (30 trials for each lever) termed "forced choice trials" (or "cued choice trials") cue lights above the levers signaled which lever, if chosen, would result in a food reward (Fig. 2.1C). Rats learned to distinguish between the two cue lights with near

perfect accuracy revealing no side bias during forced choice trials (Fig. 2.1D; $t_{(10)} = 0.305$, $p = 0.767$).

Once stable responding on this task was observed, subjects experienced one of two negative contingency switches. In one contingency switch, the reward following a correct operant response on one lever was reduced by 50% (from 2 pellets to 1 pellet); in the other contingency switch, reward was completely omitted (from 1 pellet to 0 pellets). In both cases, the reward was decreased by 1 pellet, and the unchanged lever remained reinforced as normal.

Reducing reward by 50% on one lever did not induce a behavioral preference during the first session; however, a preference for the optimal choice emerged over subsequent reward reduction sessions (Fig. 2.2A; main effect of reinforcement, $F_{(1,48)} = 86.270$, $p < 0.001$; interaction of reinforcement by session, $F_{(2,48)} = 18.082$, $p < 0.001$). Specifically, during the first session of reward reduction, subjects did not exhibit a behavioral preference for one lever over the other during free choice trials ($p = 0.545$). By the second ($p < 0.001$) and third ($p < 0.001$) sessions of reward reduction, subjects showed a significant preference for the lever yielding twice as much reward. These data demonstrate that the rats learned this contingency switch; however, a 50% reduction in reward was not a salient enough reduction to prompt an immediate alteration in behavior.

In contrast to reward reduction, when the reward following a correct operant response on one lever was unexpectedly omitted (i.e. reduced to 0 pellets), subjects displayed a robust behavioral preference for the rewarded lever during the very first session (Fig. 2.2B; main effect of reinforcement, $F_{(1,54)} = 949.129$, $p < 0.001$; interaction

of reinforcement by session, $F_{(2,54)} = 37.372$, $p < 0.001$; session 1, $p < 0.001$). A strong preference for the rewarded lever continued during free choice trials of the second ($p < 0.001$) and third ($p < 0.001$) sessions of reward omission. In fact, the preference for the rewarded lever was even stronger during the second and third sessions of reward omission ($F_{(1,12)} = 51.280$, $p < 0.001$); subjects chose the omitted reward lever significantly fewer times during the second ($p < 0.001$) and third ($p < 0.001$) sessions compared to the first session. Choice preference did not significantly increase between the second and third sessions of reward omission ($p = 0.227$) possibly due to a ceiling effect: subjects were almost exclusively choosing the rewarded lever during free choice trials already by the second session (Fig. 2.2B).

Although both reward reduction and omission spurred a preference for the more valuable option, reward omission prompted a more rapid, robust preference for the rewarded option during the very first session, while the preference for the lever yielding greater reward during reward reduction was more modest, developing over sessions (Fig. 2.2C; main effect, $F_{(1,33)} = 56.451$, $p < 0.001$). Indeed, reward omission subjects showed a significantly stronger preference for the better option lever during free choice trials than reward reduction subjects during all three contingency-switch sessions (Fig. 2.2C; session 1, $p < 0.001$; session 2, $p < 0.001$; session 2, $p < 0.001$). Reward omission was the only contingency switch which produced robust changes in behavioral preference on the first day.

No sex differences were observed in performance of the foraging task

Since sex differences exist in a variety of rodent behavioral tasks (Van Haaren et al., 1990; Jonasson, 2005; Becker and Taylor, 2008; Dalla and Shors, 2009; Sutcliffe,

2011), and should be examined in new behavioral models (Becker et al., 2005; Beery and Zucker, 2011), we included both sexes to determine if male and female rats respond differently to reduction and omission of an expected reward.

Male and female rats did not differ in baseline performance of the task (Fig. 2.3). Both male and female rats earned rewards from both levers, showing no reliable preference for one lever over the other during free choice trials, especially given that each lever was equally rewarded (Fig. 2.3A; main effect, $F_{(1,23)} = 1.389$, $p = 0.251$; males, $p = 0.819$; females, $p = 0.143$). No significant sex difference was observed in baseline performance on free choice trials (Fig. 2.3A; $F_{(1,23)} = 0.778$, p = 0.387). Both male and female subjects learned to distinguish between the two cue lights with near perfect accuracy revealing no side bias during forced choice trials (Fig. 2.3B; main effect, $F_{(1,23)} = 0.801$, p = 0.380; males, p = 0.328; females, p = 0.827), and performance between male and female rats did not significantly differ on forced choice trials (Fig. 2.3B; $F_{(1,23)} = 2.375$, p = 0.137). Furthermore, male and female rats did not significantly differ in number of sessions to meet criterion for stable responding (FR4: males mean 3.73 ± 2.41, females means: 5.00 ± 2.65, $t_{(22)} = -1.222$, p = 0.235).

Additionally, both male and female subjects displayed a preference for the lever yielding greater reward during the second and third, but not the first, session of reward reduction and a preference for the rewarded lever during all three sessions of reward omission (Fig. 2.4). Reducing reward by 50% on one lever did not initially induce a behavioral preference (main effect $F_{(1,15)} = 0.178$, $p = 0.679$). Specifically, during the first session of reward reduction neither male ($p = 1.000$) nor female ($p = 0.448$) subjects exhibited a behavioral preference for one lever over the other during free choice trials

(Fig. 2.4A), and there was no difference in choice behavior between males and females ($p = 0.679$). By the second ($F_{(1,15)} = 43.986$, $p < 0.001$) and third ($F_{(1,15)} = 39.656$, $p < 0.001$) sessions of reward reduction, male (Fig. 2.4B; session 2, $p = 0.005$; session 3, $p = 0.006$) and female (Fig. 2.4C; session 2, $p < 0.001$; session 3, $p < 0.001$) subjects showed a significant preference for the lever yielding twice as much reward. These data demonstrate that male and female rats learned this contingency switch; however, a 50% reduction in reward was not a salient enough reduction to prompt an immediate alteration in behavior.

In contrast to reward reduction, when the reward following a correct operant response on one lever was unexpectedly omitted (i.e. reduced to 0 pellets), both male ($p < 0.001$) and female ($p < 0.001$) subjects displayed a robust behavioral preference for the rewarded lever during the very first session (Fig. 2.4D; main effect $F_{(1,17)} = 55.006$, $p < 0.001$) with no sex differences ($p = 0.513$). A strong preference for the rewarded lever continued during free choice trials of the second session ($F_{(1,17)} = 475.959$, $p < 0.001$) and third session ($F_{(1,17)} = 2{,}517.712$, $p < 0.001$) of reward omission in males (Fig. 2.4E; $p < 0.001$) and females (Fig. 2.4F; $p < 0.001$).

The lack of sex difference in development of choice preference was not due to estrous cycle effects in female subjects

Estrous cycle stage for each female rat was determined via vaginal lavage (Becker et al., 2005) for all reward reduction and omission sessions. Samples containing primarily leukocytes were classified as diestrus (Fig. 2.5A). Smears containing predominantly nucleated epithelial cells were classified as proestrus (Fig. 2.5B). Smears containing predominately cornified cells were classified as estrus (Fig. 2.5C). Samples

containing a mixture of cornified and leukocytes were classified as metestrus. Metestrus and diestrus data were combined for analysis (Lynch et al., 2000), and estrous cycle stage was included as a covariate in analysis (Girard and Garland Jr, 2002; Pawluski et al., 2006) to determine if it modulated the development of the behavioral preference.

Female subjects in all stages of the estrous cycle performed similarly to males (Fig. 2.5D-I). Indeed, estrous cycle did not affect development of choice preference during the first (Fig. 2.5D; $F_{(1,14)} = 0.001$, $p = 0.973$), second (Fig. 2.5E; $F_{(1,14)} = 1.327$, $p = 0.269$) or third (Fig. 2.5F; $F_{(1,14)} = 0.015$, $p = 0.905$) session of reward reduction. Similarly, estrous cycle stage did not affect behavioral preference in the first (Fig. 2.5G; $F_{(1,16)} = 0.669$, $p = 0.425$), second (Fig. 2.5H; $F_{(1,116)} = 2.135$, $p = 0.163$), or third (Fig. 2.5I; $F_{(1,16)} = 0.002$, $p = 0.970$) session of reward omission.

Aversive components of reward omission

Reinforcement learning theory (Glimcher, 2011) would predict that through experiencing reward omission, subjects would learn to associate the cue light above the non-reinforced lever with receiving zero reward and learn that the cue light over the other lever predicts reward availability. While reinforcement learning importantly focuses on learning and the predictability of outcomes, frustration theory addresses the emotional component of reward omission, stating that the omission of an anticipated reward is aversive and "frustrating" (Amsel, 1958). Since the utilized reward omission paradigm prompts a rapid and robust preference for the rewarded lever, we hypothesized that the cues for the reward omission lever would develop aversive qualities. Specifically, we predicted that subjects would avoid the quadrant of the behavioral chamber containing the reward omission lever and reduce responding during forced choice trials on the

reward omission lever. Since reward reduction did not cause a choice preference during the first session, we did not expect that the reduced reward lever would become aversive as determined by avoidance, and that responding on forced choice trials for the smaller reward would be reduced during the first session.

Behavioral videos were analyzed to determine where rats were spending time throughout the sessions. During the inter-trial intervals when levers were unavailable (which, in total, accounted for over 50% of the length of the session), subjects could freely explore the chamber, and since levers did not become available until 5 sec after the cue light was illuminated, subjects had time to approach the lever from any place in the chamber.

Video analysis revealed that during baseline sessions when both levers were equally reinforced, subjects did not spend more time in the quadrant in front of one lever over the other (Fig. 2.6D; $t_{(16)} = -0.326$, $p = 0.748$). Similarly, subjects experiencing reward reduction did not display a significant preference for the quadrant containing the lever yielding optimal reward during the first session of reward omission (Fig. 2.6D; $t_{(14)} = 1.007$, $p = 0.331$). Supporting our hypothesis, subjects experiencing reward omission spent significantly less time in the quadrant of the chamber containing the extinguished lever and significantly more time in the quadrant containing the reinforced lever (Fig. 2.6D; $t_{(14)} = 7.084$, $p < 0.001$). These results demonstrate the salience of reward omission on goal-directed behavior and support the theory that cues signaling reward omission acquire aversive properties and therefore are avoided.

Furthermore, performance on forced choice trials during reward reduction did not significantly differ during the first session (Table 2.1). During sessions two and three, performance on forced choice trials for the smaller reward (the reward that has been reduced by 50%) modestly, but significantly, decreased (Table 2.1). Conversely, subjects experiencing reward omission showed a robust decrease in performance on omitted forced choice trials while performance on rewarded forced choice trials remained very high over all three sessions (Table 2.1). Together with the quadrant analyses and free choice results, these data demonstrate that reward omission causes greater avoidance of the suboptimal choice than reward reduction and are consistent with frustration theory showing that the omission of an expected reward is a salient and aversive event.

After three sessions of reward reduction or omission, responding on both levers was once again equally reinforced for six sessions. Although a response on either lever yielded the same reward, subjects did not initially choose both levers equally during free choice trials (main effect for performance on free choice trials: subjects that had experienced reward reduction: Fig. 2.6E; $F_{(1,5)} = 19.298$, $p < 0.001$; subjects that had experienced reward omission: Fig. 2.6F; $F_{(1,5)} = 85.810$, $p < 0.001$). Post hoc analyses revealed that subjects who previously underwent reward reduction displayed a significant preference during the first two post switch sessions for the lever that had previously signaled the larger reward (Fig. 2.6E; session 1, $p = 0.003$; session 2, $p = 0.006$), but did not display a significant preference for one lever over the other the remainder of the post switch sessions (session 3, $p = 0.097$; session 4, $p = 0.373$; session 5, $p = 0.207$; session 6, $p = 0.670$).

In contrast, subjects that had undergone reward omission maintained a significant preference for the lever that had consistently been rewarded and chose the lever that had previously resulted in an omitted reward less across all six post switch sessions (Fig. 2.6F; session 1, $p < 0.001$; session 2, $p < 0.001$; session 3, $p = 0.003$; session 4, $p = 0.024$; session 5, $p = 0.001$; session 6, $p = 0.002$). The decreased responding on the previously non-rewarded lever was not attributable to a deficit in learning (i.e. not knowing that choosing this lever would result in reward), since performance on forced choice trials for both levers was nearly perfect by the second session (Table 2.2; main effect, $F_{(1,5)} = 9.416$, $p = 0.005$). Subjects that had undergone the reward reduction contingency switch performed equivalently on forced choice trials for both levers (Table 2.2; main effect, $F_{(1,5)} = 0.889$, $p = 0.355$). These data are consistent with frustration theory (Amsel, 1958) and support the hypothesis that during reward omission the cues predicting an omitted reward are tagged with aversive motivational properties (Liu et al., 2008).

Discussion

Psychological mechanisms underlying the development of behavioral preferences mediated by aversive motivation.

It is well established that the absence of an expected reward is a salient event prompting behavioral reactions often described as emotional (Tinklepaugh, 1928; Miller and Stevenson, 1936; Crespi, 1942; Skinner, 1953; Salinas et al., 1997; Salinas and Gold, 2005; Sastre and Reilly, 2006; Young and Williams, 2010; Purgert et al., 2012; Ramot and Akirav, 2012; Veeneman et al., 2012). Based on the many observations that the omission of an anticipated reward is an aversive event that has been described as "frustrating," frustration theory emerged (Amsel, 1958; Daly, 1974). In support of this

theory, a series of studies have shown that rats will lever press or jump hurdles to escape stimuli that were previously associated with reward but now are associated with the omission of reward (Adelman and Maatsch, 1956; Daly, 1969c, b, a, 1974). Consistent with these studies, our results demonstrate that rats quickly recognize the omission of an expected reward and rapidly develop a preference for the optimal choice, avoiding the extinguished lever and the quadrant of the behavioral chamber containing that lever.

A 50% reduction of reward (that was an equivalent decrease in the number of pellets as the omission condition; i.e. one less reward pellet) eventually elicited a similar behavioral preference but not as quickly or robustly as reward omission. Consistent with previous studies (Salinas et al., 1993; Salinas and White, 1998; Sastre and Reilly, 2006; Ramot and Akirav, 2012), reward reduction evoked a significant preference for the more valuable option by the second session (24 hrs later). Multiple studies that have reduced reward value by 90% have observed behavioral effects during the first session including increased latency to retrieve reward in a maze (Salinas et al., 1996; Salinas et al., 1997; Kerfoot et al., 2008) and consuming less of the reward (Salinas and Gold, 2005). Together with our results, these findings indicate that reward reduction can have significant, immediate effects on behavior, but the reduction must be highly salient to the animal, often at levels close to omission.

Lack of sex differences in behavioral performance during the negative contingency switches

Female rats showed similar behavior as male rats in both reward reduction and omission tests, and estrous cycle stage did not affect choice preference during either contingency switch. These results are consistent with foraging studies in the field (Clark,

1980). Similarly, other laboratory studies have neither found sex differences during certain operant tasks (Van Haaren et al., 1987; Van Hest et al., 1988, 1989; Stratmann and Craft, 1997; Carroll et al., 2009) nor estrous cycle effects on operant performance (Stratmann and Craft, 1997; Davis et al., 2008; Cummings et al., 2011).

Although sex differences exist in many motivated behaviors, the lack of sex differences in this paradigm is likely due to the fact that basic flexibility in foraging strategy is adaptive in both males and females. However, although no sex differences in this behavioral paradigm exist in adult virgin rats, females caring for offspring may develop a preference for the optimal choice more rapidly than males and nulliparous females, since maternal females must forage for their pups in addition to themselves (Kinsley et al., 1999; Love et al., 2005).

Conclusion

The present experiments establish a laboratory model of foraging behavior in rats (Porter-Stransky et al., 2013). This behavioral paradigm can capture changes in choice behavior following manipulations of reward availability. Specifically, the omission of an expected reward in this behavioral task prompts a strong preference for the rewarded option in both male and female rats and across all stages of the estrus cycle. This behavioral paradigm provides a model for our future experiments to elucidate the neural mechanisms that mediate changes in reward-seeking behavior due to the omission of an expected reward (Porter-Stransky et al., 2013). Since reward omission causes the most robust choice preference during the first session, this is the manipulation that we will utilize in the remaining chapters.

Figures

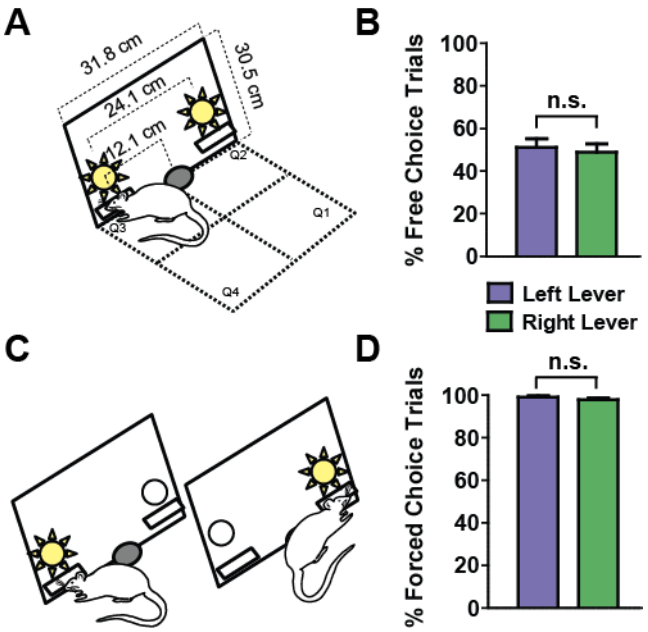

Figure 2.1. Appetitive operant behavioral paradigm for examining foraging preference in rats. **A)** Free choice trials facilitate the assessment of an animal's preference for one lever over the other. During training, choosing either lever resulted in equal amount of food reward. **B)** Once trained on the task, rats accurately completed free choice trials, showing no reliable side bias. **C)** During forced choice trials, although both levers are extended, subjects only received a food reward for pressing the lever under the illuminated cue light. **D)** Subjects learned to complete forced choice trials with near perfect accuracy. n.s. = not statistically significant, Error bars indicate mean ± SEM.

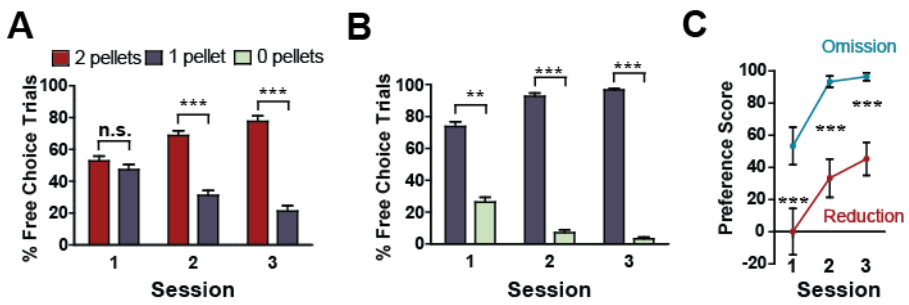

Figure 2.2. Effects of unexpected reward reduction and reward omission on choice preference. **A)** During the first session of reward reduction when the reward resulting from a correct operant response on one lever was decreased by 50%, subjects displayed no preference for the lever yielding twice as much reward. By the second and third sessions of reward reduction, rats exhibited a preference for the lever yielding greater reward during free choice trials. **B)** When the reward resulting from a correct operant response on one lever was unexpectedly decreased by 100%, a robust preference for the rewarded lever was observed that continued during all 3 sessions of reward omission. **C)** Preference for the better option during free choice trials was significantly stronger for subjects experiencing reward omission than for those experience reward reduction. n.s. = not statistically significant, $^{**}p < 0.01$, $^{***}p < 0.001$. Error bars indicate mean ± SEM.

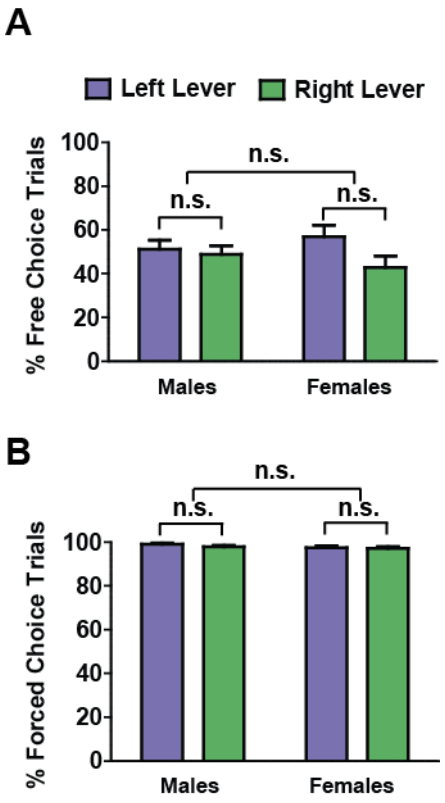

Figure 2.3: Male and female rats performed the behavioral task equally well. **A)** Neither male nor female rats exhibited a significant preference for one lever or the other. **B)** Both males and females completed forced choice trials with near perfect accuracy. n.s. = not statistically significant. Error bars indicate mean ± SEM.

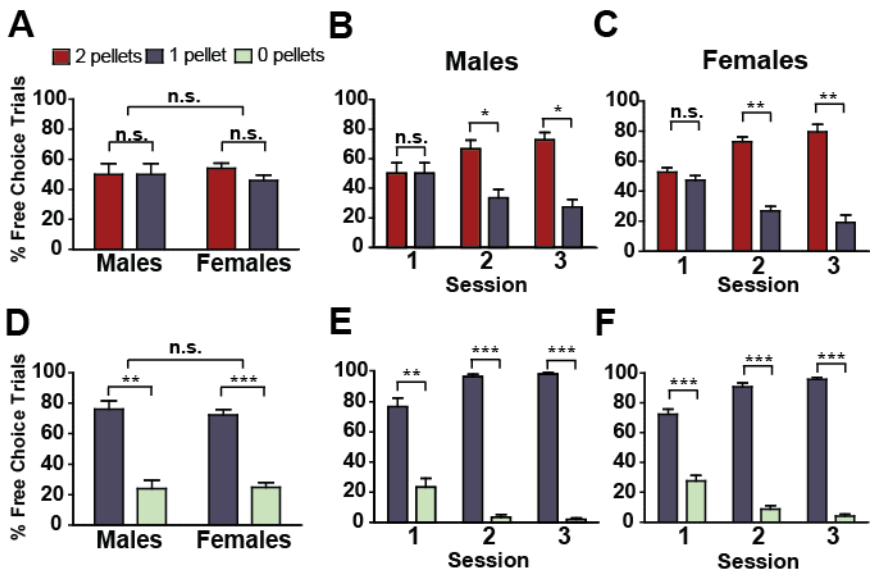

Figure 2.4: Male and female rats exhibited the same choice preferences during the two negative contingency switches. **A-C**) Neither male nor female rats displayed a choice preference during the first session of reward reduction (A). Both male (B) and female (C) rats displayed a preference for the lever yielding greater reward during the second and third sessions of reward reduction. **D-E)** Both males and females exhibited a significant preference for the rewarded lever during the first reward omission session (D). Males (E) and females (**F**) similarly showed a robust preference for the rewarded option during all sessions of reward omission. n.s. = not statistically significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Error bars indicate mean ± SEM.

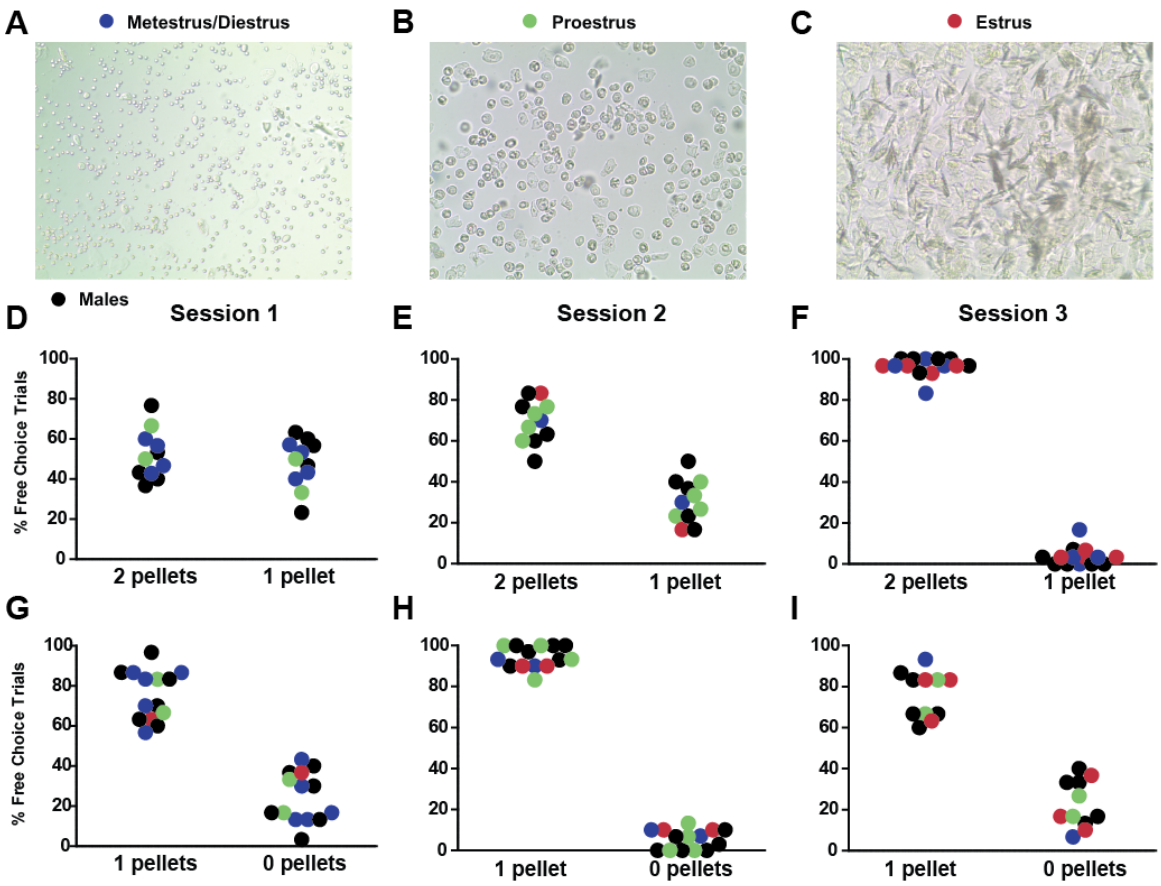

Figure 2.5. Estrous cycle did not affect the development of choice preference. **A-C)** Representative images of vaginal epithelial cells in diestrus (A), proestrus (B), and estrus (C). **D-F)** Female rats in each estrous cycle stage performed similarly to males (black circles) on the first (D), second (E), and third (F) session of reward reduction. **G-I)** During the first (G), second (H), and third (I) session of reward omission, females in each estrous cycle stage performed similarly to males.

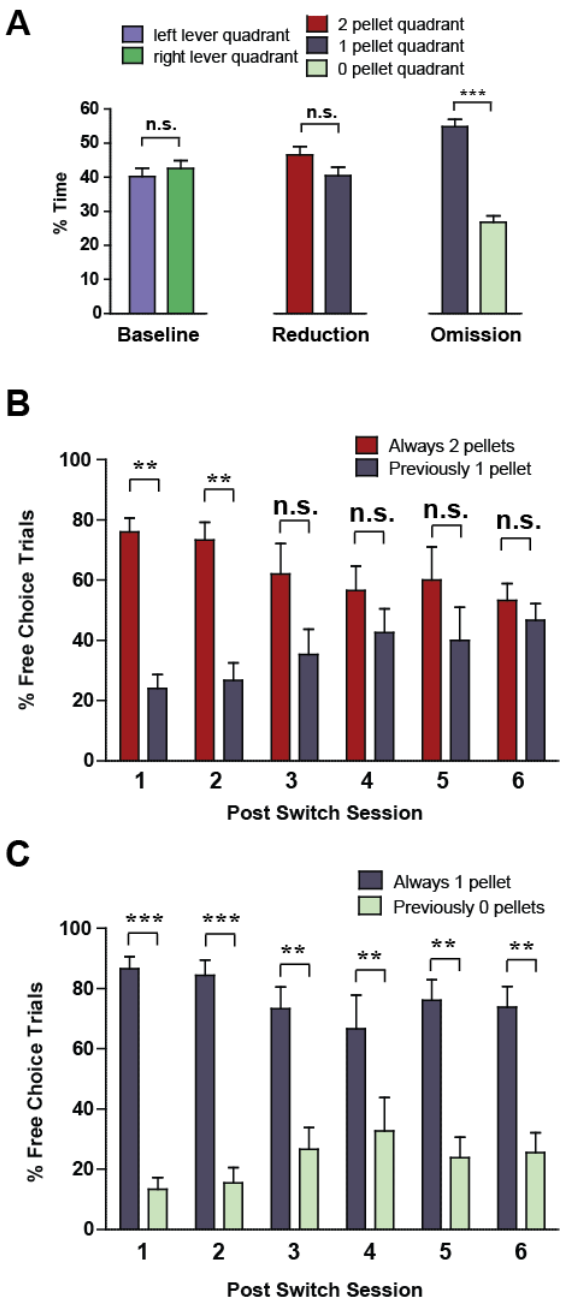

Figure 2.6. Additional behavioral effects of reward omission. **A**) Percentage of time spent in the quadrants containing of the levers did not differ during baseline sessions or the first session of reward reduction; however, during the first session of reward omission, rats spent significantly more time in the quadrant containing the rewarded lever than the quadrant containing the non-reinforced lever. **B-C**) When the levers were once again equally reinforced, subjects that had experienced reward reduction (B) lost the behavioral preference by the third session, whereas subjects that had experienced reward omission

(**C**) maintained a preference for the lever that continually had been reinforced. n.s. = not statistically significant, $^{**}p < 0.01$, $^{***}p < 0.001$. Error bars indicate mean ± SEM.

Table 2.1. Percentage of forced choice trials accurately completed.

	Reward Reduction		Reward Omission	
	2 Pellets	1 Pellet	1 Pellet	0 Pellets
Session 1	96.47 ± 1.32	94.29 ± 1.71	97.72 ± 0.88	71.57 ± 4.58
	$F_{(1,15)} = 0.901$, $p = 0.358$		$F_{(1,17)} = 24.849$, $p < 0.001$	
Session 2	99.216 ± 0.45	93.33 ± 1.43	98.42 ± 0.78	59.82 ± 6.16
	$F_{(1,15)} = 8.555$, $p = 0.010$		$F_{(1,17)} = 28.085$, $p < 0.001$	
Session 3	98.24 ± 0.91	85.69 ± 2.26	98.42 ± 0.74	27.19 ± 4.53
	$F_{(1,15)} = 22.594$, $p < 0.001$		$F_{(1,17)} = 208.197$, $p < 0.001$	

Table 2.2. Percentage of post switch forced choice trials accurately completed. During these trials, responses on both levers were once again equally rewarded.

	Post Reward Reduction		Post Reward Omission	
	Always 2 Pellets	Previously 1 Pellet	Always 1 Pellet	Previously 0 Pellets
Session 1	98.67 ± 0.82	98.00 ± 1.33	99.44 ± 0.56	77.78 ± 8.59
	$p = 0.569$		$p < 0.001$	
Session 2	99.33 ± 0.67	99.33 ± 0.67	100.00 ± 0.00	98.33 ± 0.75
	$p = 0.260$		$p = 0.635$	
Session 3	98.00 ± 1.33	98.67 ± 0.82	100.00 ± 0.00	97.22 ± 1.34
	$p = 0.569$		$p = 0.430$	
Session 4	98.00 ± 1.33	98.00 ± 1.33	99.44 ± 0.56	100.00 ± 0.00
	$p = 1.000$		$p = 0.874$	
Session 5	99.33 ± 0.67	98.67 ± 1.33	99.44 ± 0.56	99.44 ± 0.56
	$p = 0.569$		$p = 1.000$	
Session 6	99.33 ± 0.67	98.67 ± 0.82	100.00 ± 0.00	99.44 ± 0.56
	$p = 0.569$		$p = 0.874$	

CHAPTER 3:

INCREASES IN EXTRACELLULAR DOPAMINE IN THE NUCLEUS ACCUMBENS DURING REWARD OMISSION AND CORRESPONDING INCREASES IN MOTIVATIONAL VIGOR AND EXPLORATION OF ALTERNATIVE RESPONSE STRATEGIES

Introduction

The ability to recognize changes in resource availability and alter behavioral strategy accordingly is fundamental to adaptive, reward seeking behavior. Dopamine (DA) transmission has been proposed to modulate behavior flexibility rather than just being a "reward neurotransmitter" (Beeler et al., 2014). Consistent with this idea, pharmacological manipulations affecting endogenous DA transmission have been shown to impair or enhance certain forms of behavioral flexibility (Ragozzino, 2002; Frank, 2005; Floresco et al., 2006a; Floresco and Magyar, 2006; Floresco et al., 2006b; Haluk and Floresco, 2009; Winter et al., 2009; Skelin et al., 2014). Recently, we have shown that rats promptly develop a preference for a rewarded option when another option that previously was reinforced is extinguished, and this form of behavioral flexibility is mediated by D2-like DA receptors in the nucleus accumbens (NAc) (Porter-Stransky et al., 2013). However, the DA transmission dynamics that may facilitate this behavior have yet to be elucidated.

Unlike many other neurotransmitters which are confined to the synaptic cleft when released, DA can escape the synaptic cleft and signal through volume transmission (Arbuthnott and Wickens, 2007) and on multiple time scales (Grace et al., 2007; Baudonnat et al., 2013). Transient, sub-second surges in DA release are thought to be the result of burst firing of midbrain DA neurons, and much attention has been given to these phasic release events since they frequently occur in relation to behaviorally relevant events (Schultz, 2002; Day et al., 2010; Gan et al., 2010). In addition to burst firing (Grace and Bunney, 1984a), DA neurons also exhibit a low, single-spike firing rate (Grace and Bunney, 1984b). This baseline firing rate of DA neurons, as well as effects from terminal modulation (Grace, 1991; Keefe et al., 1993; Howland et al., 2002), are likely responsible for tonic, extra-cellular levels of DA concentration ([DA]) in striatal regions that are captured by microdialysis (Floresco et al., 2003; Grace et al., 2007); however, phasic release of DA also may contribute to these levels (Owesson-White et al., 2012).

While phasic DA activity has been proposed to mediate temporally precise learning about changes in reward availability (Schultz et al., 1997; Steinberg et al., 2013) or the attribution of incentive motivational properties to cues (Berridge, 2012), tonic levels of striatal DA have been hypothesized to signal other important factors contributing to motivated behavior that occur over longer time scales. For example, tonic levels of DA have been suggested to track the average rate of reward (Niv et al., 2007), the average rate of punishment (Daw et al., 2002; Daw and Touretzky, 2002), the vigor of behavioral responses (Niv et al., 2005; Niv, 2007), effort (Salamone et al., 2007), thrift (Beeler, 2012), and the balance between exploration and exploitation (Beeler et al., 2010;

Humphries and Prescott, 2010). Although there appear to be numerous competing theories for the role of tonic levels of DA, there are common themes among them. First, DA may track the goodness of the available options; this could encompass theories of DA's role in signaling reward rate (Niv, 2007) or punishment rate (Daw et al., 2002). Second, DA could signal motivational vigor, which could include response vigor (Niv et al., 2007) and incentive salience (Berridge and Robinson, 1998). Third, DA may signal changes in behavioral strategies, which could include the trade-off between explorative versus exploitive strategies (Humphries et al., 2012; Beeler et al., 2014). Additionally, this could include the regulation of energy expenditure, including thrift (Beeler, 2012), and response strategies based on effort (Salamone et al., 2007; Salamone and Correa, 2012). All of these factors could importantly contribute to behavioral flexibility when reward outcomes change.

One of the primary hurdles in determining the relationship between tonic levels of DA and the aforementioned behaviors has been the lack of a suitable technique for capturing such changes in DA. Electrophysiology is excellent for capturing changes in baseline firing rate of neurons; however, there have been issues with reliable identification of VTA DA neurons *in vivo* (Margolis et al., 2006a; Ungless and Grace, 2012), and electrophysiology does not capture other sources of neurotransmission that affect tonic levels of DA in the striatum. While fast-scan cyclic voltammetry (FSCV) provides sub-second temporal resolution and excellent spatial resolution, it is inherently a differential technique, limited to examining change in [DA] over 30-90 sec periods of time. Since a large current is background subtracted out, FSCV cannot examine absolute levels of [DA] over minutes to hours. Microdialysis, although having inferior temporal

and spatial resolution compared to FSCV, is able to examine changes in [DA] over minutes to hours. Traditionally, microdialysis conducted in behaving animals generates one sample every 10-20 min. Since numerous behaviors can occur within these long ranges of time, the precise relationship between DA and behavior has been obscured. Recent advances in analytical chemistry have significantly improved the temporal resolution of microdialysis if samples are analyzed with mass spectrometry (Song et al., 2012) instead of traditional electrochemical detection.

Here, we utilize a one minute sampling microdialysis in the NAc of rats engaged in a task prompting changes in choice behavior due to the omission of an expected reward (Porter-Stransky et al., 2013). In addition to examining changes in [DA] during the extinction of one lever, we also examine the relationship between changes in [DA] and a number of behavioral measures to test some of the hypotheses of tonic DA transmission. If [DA] tracks the "goodness" of the current options, either by signaling average rate of reward (Niv et al., 2007) or rate of punishment (Daw et al., 2002), [DA] would be expected to decrease during periods of reward omission and increase when reward is available, or vice versa. If DA signals behavioral motivation (Niv, 2007), then [DA] could correspond to the vigor of behavioral responding. Finally, if tonic levels of DA regulate behavioral strategy, including thrift (Beeler, 2012) and the balance between exploitation and exploration (Beeler, 2012; Humphries et al., 2012), then a relationship between the allocation of these behaviors and [DA] could be expected.

Methods

Subjects

Data from a total of 13 male, Sprague Dawley rats, obtained from Charles River (Winington, MA), were included in this experiment (experimental group, n = 8; control group, n = 5). Animals were kept on a reverse light/dark cycle, and tested during their dark cycle. Subjects had free access to water in their cages and were mildly food restricted to approximately 90% of free-feeding weight, as previously described (Porter-Stransky et al., 2013).

Behavioral Paradigm

The operant behavioral task to model foraging behavior was the same as previously described in chapter 2 (Porter-Stransky et al., 2013). The task included two trial types, cued choice trials (previously termed forced choice trials) and free choice trials (Fig. 3.1). Both trial types were interspersed throughout the session. Cued choice trials were important to ensure that subjects were discriminating between the two cue lights, and free choice trials probed for preferences for one lever over the other. Rewards were 45 mg BioServ chocolate pellets.

On cued choice trials, one cue light would illuminate, and then five seconds later, both levers would extend. Subjects could earn a reward by choosing the lever under the illuminated cue light (Fig. 3.1A). Pressing the lever under the non-illuminated cue light was considered an error and not rewarded. After a time out, the next trial began. On free choice trials, both cue lights were illuminated, and subjects could earn a reward on either

lever. Throughout training, the illuminated levers were always rewarded (Fig. 3.1A-B). Subjects were trained on the FR4 version of the task (Porter-Stransky et al., 2013).

During microdialysis and FSCV test sessions, subjects first received 18 trials of this baseline version of the task. Then the reward omission manipulation began, during which the reward usually given from a correct response on one lever was omitted. The other lever continued to be reinforced as usual (Fig. 3.1C-D). The non-rewarded lever remained consistent within a subject for the duration of the test session in both cued choice trials (Fig. 3.1C) and free choice trials (Fig. 3.1D); however, which lever was extinguished was counter-balanced across subjects. Throughout training, the trial types were pseudo-randomly interspersed. Since the rate of collecting dialysis fractions was 1 min, trial types were blocked during the test session so that each min only contained one trial type. Throughout the results sections in chapters 3 and 4, trials during the reward omission session are coded as follows: cued choice trials for the rewarded lever are coded in green, cued choice trials for the non-rewarded lever are coded in red, and free choice trials are coded in blue. Each of these trial types prior to experiencing reward omission (i.e. baseline and the control group) are coded in gold.

Behavioral Analysis

MED Associates (St. Albans, VT) software recorded which lever subjects pressed each trial as well as the duration of time from lever extension until subjects completed lever pressing each trial. In addition to lever pressing, subjects displayed a number of other behaviors during the trials and inter-trial intervals that could serve as proxies for exploratory behavior. Specifically, using Behavior Tracker software, we coded rearing behavior, grooming behavior, time exploring the back half of the chamber (where no cues

or food was available), and when subjects engaged the cues (including the cue lights and the ports where the levers recess into the wall) and food cups, defined as orienting toward, sniffing, biting, or pawing at these stimuli. We documented each occurrence of these behaviors and also calculated the number of times subjects switched from one event to another during each trial type within each session block.

Surgery

All procedures were approved by the University of Michigan Committee on the Use and Care of Animals and performed in accordance with their policies. Subjects were anesthetized with an intramuscular injection of ketamine hydrochloride (90 mg/kg) and xylazine hydrochloride (10 mg/kg). Rats were also given the analgesic ketoprofin (5 mg/kg) and antibiotic cefazolin (70 mg/kg).

During surgery, guide cannulae (Eicom, San Diego, CA) were positioned over the NAc core (AP: +1.4; ML: +/- 1.3) and fixed in place with skull screws and dental acrylic. Additionally, a flattened wound clip was cemented onto the headcap to tether subjects during the microdialysis experiment. Subjects were fed *ad libitum* until fully recovered from surgery, after which food restriction and behavioral training resumed.

After the experiment had concluded, subjects were euthanized with sodium pentobarbital (70 mg/kg). Brains were sliced on a cryostat at 50 µm, and probe placements were examined under 10x magnification on a microscope.

Microdialysis

On the day of testing, microdialysis probes were inserted into the guide cannulae. Probes projected 1mm below the cannulae to sample from fresh tissue. Artificial

cerebrospinal fluid (aCSF) was perfused through the probe at the rate of 2 µL/min using a Chemyx Fusion pump (Stafford, TX). Consistent with previous experiments (Colivicchi et al., 2013; Hossain et al., 2013), we waited at least 90 min after the probe was inserted into brain before dialysate collection began. For the duration of the experiment, fractions were collected every 60 sec.

To establish baseline levels of [DA] prior to the behavioral task, samples were collected for 9 min. Then the behavioral task began. Trials were blocked so that each dialysate sample contained only one trial type. Three trials occurred each min. Subjects were first given 18 trials of the baseline operant task in which both levers were equally reinforced (Fig.3.1). Then, the reward omission manipulation began in which the reward following a correct response on one lever was omitted while the other lever continued to be reinforced as usual (Fig. 3.1).

Throughout the experiment, each 2 µL sample was immediately derivatized as previously described (Song et al., 2012). Specifically, to prepare fractions for analysis, 1.5 µL of 100mM sodium tetraborate, 1.5 µL of benzoyl chloride in 2% acetonitrile, and 1.5 µL of stable-isotope labeled internal standards were added, in that order, to each sample of dialysate.

Fractions were analyzed using high performance liquid chromatography with mass spectrometry via a nanoAcquity HPLC system (Waters, Milford, MA) equipped with a Waters 1 mm x 100 mm HSS T3 reverse-phase HPLC column operated at 100 µL/min. Eluting analytes were detected using an Agilent 6410 triple quadrupole MS

(Agilent, Santa Clara, CA) operating in positive mode performing dynamic multiple-reaction-monitoring. The limits of detection for DA on this instrument were 0.0969 nM.

Statistics

Statistical analyses were performed using IBM SPSS Statistics 21 (Armonk, NY) and GraphPad Prism (La Jolla, CA). To examine minute by minute changes in [DA], a linear mixed model was used for its ability to handle correlated, repeated measures data, as previously described (Aragona et al., 2008; Aragona et al., 2009). This model was used to examine within-subject changes in [DA] as well as to compare control subjects to the experimental group receiving reward omission (Fig. 3.2A). The model adjusts degrees of freedom based on the distance between comparisons, which is why reported degrees of freedom values are often non-integers.

Changes in choice preference as well as the other behavioral measures were analyzed using repeated measures ANOVAs. Dunnett's post hoc tests were used to examine changes in specific behaviors during the different trial types in the reward omission session compared to baseline levels of those behaviors (Fig. 3.3A-C,E,G,I). Bonferroni corrections were used for the remainder of statistical tests in which multiple comparisons were performed (Fig. 3.2B-D and 3.3D). Finally, linear regression was used to test for correlations between [DA] and behavior.

Results

To examine changes in [DA] across minutes in response to reward omission, one min sampling microdialysis was employed. During the baseline trials when both levers were equally reinforced (Fig. 3.1A-B), [DA] significantly increased to ~150% of basal

levels (Fig. 3.2A, main effect, $F_{(6, 54.572)} = 2.619$, $p = 0.027$). During these trials, subjects sampled from both levers equally (Fig. 3.2B) and there were no differences in [DA] among trial types ($F_{(2, 36)} = 0.169$, $p = 0.845$).

When the reward omission manipulation began, [DA] increased to ~200% of basal levels. Indeed, [DA] significantly increased during the very first trials of reward omission (min 16, $p = 0.028$; min 17, $p = 0.044$; min 18, $p = 0.054$), and [DA] was significantly greater in subjects experiencing reward omission compared to subjects performing the regular task with both levers being equally reinforced (Fig. 3.2A; effect of group in blocks 1-3, $F_{(1, 46.059)} = 16.324$, $p < 0.001$). The elevation in [DA] during reward omission trials persisted throughout the entire session (Fig. 3.2A). Importantly, changes in [DA] in the reward omission group were not attributable to any inherent differences between the two groups. Indeed, raw levels of basal [DA] did not differ between the two groups prior to the behavioral experiment ($t_{(11)} = 1.821$, $p = 0.100$; experimental group mean, 0.683 nM; control group mean, 0.195 nM). Additionally, percentage of basal levels across the 9 fractions prior to the behavioral experiment did not vary between groups, $F_{(1, 44.631)} = 0.459$, $p = 0.502$) or during the baseline task prior to the reward omission switch ($F_{(1, 16.785)} = 2.292$, $p = 0.149$).

Behaviorally, subjects quickly developed a preference for the rewarded lever during free choice trials (Fig. 3.2C; interaction of reinforcement by block, $F_{(2,21)} = 4.167$, $p = 0.030$) that was evident during the very first block ($p < 0.001$) and continued for the remainder of the session ($p < 0.001$ on blocks 2 & 3 as well, Bonferroni corrections). Additionally, subjects reduced responding on cued choice trials for the non-rewarded lever while continuing to respond correctly on the rewarded lever (Fig. 3.2D; interaction

of reinforcement by block, $F_{(2,21)} = 8.925$, $p = 0.002$). However, the reduced responding during non-rewarded cued choice trials did not emerge until the second block (post hoc tests with Bonferroni corrections: Block 1, $p = 0.664$; Block 2, $p = 0.075$; Block 3, $p < 0.001$).

Interestingly, after experiencing reward omission, [DA] remained elevated for a number of the rewarded trials as well (Fig. 3.2A). Indeed, while [DA] increased following reward omission, levels of [DA] did not correlate with number of rewards earned (Fig. 3.2E), suggesting that tonic levels of DA in the NAc do not signal reward rate.

Since tonic levels of [DA] have been proposed to signal the vigor of behavioral response (Niv et al., 2007) as well as the balance between exploring new options and exploiting known resources (Beeler et al., 2012), we examined changes in response vigor and exploratory behaviors throughout the session. Response vigor was determined by the amount of time it took subjects to complete the required lever presses on the lever underneath the illuminated cue light for each trial. Three measures of exploratory behavior were quantified: 1) amount of time spent investigating the back of the behavioral chamber away from the lever; 2) percentage of time rearing; and 3) frequency in which rats switched their behavioral focus (such as moving from one lever to the other lever, running to the back of the chamber, grooming, etc.). Finally, behavioral responses on the two levers were used as a proxy for exploitative behavior.

Consistent with the vigor hypothesis of DA, as [DA] increased during the first few trials of reward omission (Fig. 3.1F; $t_{(6)} = 3.060$, $p = 0.018$), latency to complete the

ratio requirement decreased on the non-rewarded lever (Fig. 3.2F; $t_{(6)} = 3.199$, $p = 0.019$). Indeed, there was a significant correlation between response vigor and [DA] across trial types (Fig. 3.2G; linear regression, $R^2 = 0.175$, $p = 0.001$), indicating that subjects lever pressed faster when levels of [DA] were higher in the NAc.

In addition to responding more vigorously, rats also displayed a number of behaviors indicative of exploitation during periods of reward availability and exploration when no reward was available. Subjects spent significantly more time engaging the reward predictive cues during trials in which they could obtain reward compared to baseline levels of engaging those same cues (Fig. 3.3A; main effect, $F_{(9, 63)} = 9.956$, $p < 0.001$; Dunnett's post hoc, $p < 0.05$ during cued rewarded trials in each block and free choice trials in blocks 2 &3). Conversely, during trials in which subjects could obtain reward, they spent significantly less time engaging the cues signaling no reward availability (Fig. 3.3B; main effect, $F_{(9, 63)} = 14.763$, $p < 0.001$; Dunnett's post hoc, $p < 0.05$ during cued rewarded trials in each block and free choice trials in blocks 2 &3). Additionally, subjects decreased the amount of time engaging the cues during cued trials for the non-rewarded lever throughout the session (Fig. 3.3B inset; main effect, $F_{(2,14)} = 6.632$, $p = 0.009$.). Subjects spent less time engaging the non-rewarded cues specifically during non-reinforced cued choice trials within block 2 ($p = 0.009$) and block 3 ($p = 0.017$) compared to the first block. And the linear contrast was significant (Fig. 3.3B inset, $F_{(1, 7)} = 9.790$, $p = 0.017$), indicating that subjects decreased engagement with the non-rewarded cues in a linear fashion throughout the session. To facilitate easier visual comparisons of behaviors with levels of DA, changes in [DA] from Fig. 3.2A are shown in 3 min bins in Fig. 3.3C.

In addition to spending less time engaging the non-rewarded cues, subjects employed alternative response strategies, including pressing the non-illuminated lever and not pressing either lever. These alternative response strategies occurred significantly more often on cued-choice trials for the non-rewarded lever than on cued-choice trials for the rewarded lever (Fig.3.3D; main effect of trial type, $F_{(1,21)} = 19.353$, $p < 0.001$; interaction of trial type by block, $F_{(2,21)} = 8.524$, $p = 0.002$). This effect was predominantly seen during the second and third blocks (Fig. 3.3D; Block 2, $p = 0.024$; Block 3, $p < 0.001$; post hoc tests with Bonferroni adjustments).

During trials in which no reward could be obtained, subjects spent a significantly greater amount of time engaged in other exploratory behaviors. Since exploratory behavior has been hypothesized to correspond to tonic levels of [DA] (Beeler et al., 2012; Humphries et al., 2012), we also correlated changes in [DA] to exploratory behaviors. Rats spent more time rearing during each block of reward omission trials compared to baseline levels of rearing (Fig. 3.3E; main effect, $F_{(9, 63)} = 4.021$, $p < 0.001$; Dunnett's post hoc, $p < 0.05$ during reward omission trials of each block), and there was a significant positive correlation between percentage of time rearing and [DA] (Fig. 3.3F; linear regression; $R^2 = 0.074$, $p = 0.043$). Additionally, consistent with the exploratory hypothesis, subjects switched their behavioral focus significantly more times during reward omission trials in each block (Fig. 3.3G; main effect, $F_{(9, 63)} = 7.617$, $p < 0.001$; Dunnett's post hoc, $p < 0.01$ during reward omission trials of each block) but not during the other trial types ($p > 0.05$), and there was a significant positive correlation between the number of times in which subjects switched their behavioral focus and [DA] (Fig. 3.3H; linear regression, $R^2 = 0.100$, $p = 0.003$). Finally, subjects explored the back half

of the chamber away from the task-related cues significantly longer during the non-rewarded trials (Fig. 3.3I; $F_{(9, 63)} = 4.447$, $p < 0.001$) during all three blocks (Block 1, $p = 0.020$; Block 2, $p = 0.076$; Block 3, $p = 0.014$), and amount of time exploring the back of the chamber significantly correlated with increases in [DA] across trial types (Fig. 3.3J; $R^2 = 0.119$, $p = 0.001$).

Discussion

Tonic changes in [DA] during the reward omission task

Using one-minute sampling microdialysis coupled with high performance liquid chromatography-tandem mass spectrometry, we examined changes in extra-cellular levels of DA in the nucleus accumbens (NAc) throughout the instrumental foraging task. We found that during the baseline, rewarded version of the task, DA concentration ([DA]) increased to ~150% of basal levels. Then, when the reward omission manipulation began, [DA] increased to ~200% basal levels. Interestingly, [DA] remained elevated during many of the rewarded trials as well as the non-rewarded trials.

These results support a prediction in a temporal difference algorithm made by Daw that, under situations of negative reward or aversion, a prolonged increase in tonic levels of DA was expected to occur (Daw et al., 2002; Daw and Touretzky, 2002). Tonic levels of DA were suggested to represent the average rate of punishment (Daw et al., 2002). The increase we observe in [DA] following the reward omission manipulation is consistent with this view; however, it does not appear consistent with why [DA] is higher during the baseline, rewarded version of the task (prior to the omission manipulation) compared to pre-task levels of [DA].

The increase in [DA] following reward omission is also consistent with other studies revealing an increase in [DA] during the extinction of a reward (Ahn and Phillips, 2007). It is worth noting that some experiments have shown no changes in [DA] during extinction (Lecca et al., 2006), while others have shown a decrease in [DA] (Ranaldi et al., 1999). Differences in [DA] during periods of non-reward across these studies are likely attributable to the marked differences in the experimental designs, differences in the types of reward being extinguished (food versus different drugs of abuse), or to differences in the frequency of collecting dialysis fractions. Additionally, most previous studies were constrained to substantially longer sampling periods (10-20 min per sample), which could obscure dynamic changes in [DA] across minutes (Kennedy, 2013).

DA and Behavioral Vigor

Another hypothesis in the reinforcement learning field is that tonic levels of [DA] track the average rate of rewards and thereby determine the vigor of responding (Niv, 2007; Niv et al., 2007). Our results, which show that DA levels remain above baseline during the majority of non-rewarded and rewarded trials following the reward omission manipulation, initially appear incongruent to this hypothesis, since there is explicitly not a relationship between reward rate and tonic levels of DA. However, an integral component of Niv and colleagues' theory is that response rates are optimized by balancing the desire to obtain rewards quickly and the costs of effort and time. Importantly, slow responding delays reward acquisition; therefore, sluggish responding results in the cost of wasted time (Choi et al., 2014). Therefore, in situations when reward rate is high, it can be advantageous for hungry subjects to vigorously respond to obtain as much reward as possible, since slower responding accrues the cost of time spent

not earning rewards (Niv et al., 2005). Just as the "cost of sloth" can be detrimental when one is hungry and potential reward rate is high (Niv et al., 2007), lazy behavior can be counter-productive when hungry and resources are low. So when reward is unexpectedly omitted, invigorated behavioral responding may be advantageous to energize the hungry animal to examine what changes have occurred in the environment and to explore other potential avenues of acquiring reward.

Consistent with this logic, our data reveal that as [DA] increases during reward omission, response vigor also increases. In fact, there is a linear relationship between response vigor and [DA] throughout the test session. Subjects completed the ratio requirement of lever pressing faster when DA levels were higher. Therefore, while under some circumstances, response vigor may be greater during periods of high reward (Niv et al., 2007), this does not preclude the possibility that response vigor could increase during other salient environmental factors, such as the omission of an expected reward. In fact, reacting to diminished resource availability or aversive situations might require just as much, if not even higher, levels of invigorated behavior. Supporting this notion, other studies have observed increased DA in the NAc during aversive situations (Louilot et al., 1986; Horvitz, 2000; Levita et al., 2002; Badrinarayan et al., 2012). Interestingly, the increases in [DA] during aversive conditions were higher when animals had the opportunity to escape the situation (Rada et al., 1998). Additionally, depletions of NAc DA (McCullough et al., 1993) or the administration of DA receptor antagonists (Prinssen et al., 1996; Natesan et al., 2006) are sufficient to attenuate escape behavior, suggesting a necessary role of DA in the NAc for active avoidance. Finally, research utilizing viral

gene therapy with DA deficient mice has shown that striatal DA is necessary for both the acquisition and maintenance of instrumental shock avoidance (Darvas et al., 2011).

With the wealth of data revealing increased DA transmission in the NAc during aversive events that correspond to increases in behavioral responses, DA cannot merely be tracking reward rate or mediating only appetitive behaviors. Perhaps the most parsimonious explanation consistent with the aforementioned studies is that tonic, extra-cellular levels of DA in the ventral striatum may scale the appropriate motivational vigor of behavior, regardless of whether the valence is appetitive or aversive. A number of factors including motivational state, reward availability, potential predation, and other dynamics of the current environment would likely contribute to tonic levels of DA in the ventral striatum and motivational vigor. This interpretation is consistent with the "opportunity cost of time" theory (Niv, 2007), since when the cost of slow responding is too high, the solution is more vigorous responding (Niv, 2007). Therefore, it is plausible that, regardless of valence of the situation, tonic DA levels in the NAc may signal the vigor of behavioral actions, which could be higher during appetitive or aversive situations.

DA and the Balance of Exploration versus Exploitation

A fundamental issue in decision making is determining what the best course of action is for the current state of the environment. This issue holds true for a wide range of circumstances and is applicable to human and non-human animals (March, 1991; Cohen et al., 2007; Fang and Levinthal, 2009). In a foraging environment, individuals do not always have complete knowledge of all available options; therefore, they need to explore the environment. While spending time investigating the environment is necessary, there is the cost of time and potentially wasted opportunities of gathering

resources (Beeler et al., 2012). Animals have to balance the amount of time spent exploring the environment to learn about the options versus the amount of time actively obtaining and consuming resources. Even after an animal has initially explored its options, resource availability can unexpectedly change. Therefore, individuals face the decision of continuing a strategy that is currently working or exploring alternative options, which may be superior or may be worse. This trade-off has been termed the exploration-exploitation dilemma (Laureiro-Martínez et al., 2010). Strongly favoring either an exploration-focused strategy or an exploitation-dominant can lead to suboptimal results (March, 1991; Laureiro-Martínez et al., 2010); a balance between the two is important, and which strategy is optimal depends on the environmental state.

Generally, if the current environment is stable and one's strategy is yielding plenty of reward, exploiting knowledge of the current environment and strategy can be the optimal choice. However, if conditions are changing, exploring may be more advantageous to gain new information and discover other possible outcomes (Humphries et al., 2012). This framework poses clear predictions for subjects' behavior in the present study. Initially, since subjects have undergone many days of training for the task, errors during baseline performance of the cued choice trials should be minimal. However, when environmental contingencies change following the reward omission manipulation, subjects would be expected to attempt alternative response strategies and increase exploratory behavior.

Consistent with these hypotheses, subjects accurately perform the baseline version of the task making very few errors on cued choice trials and sampling from both levers during free choice trials. Critically, following the reward omission manipulation, rats

exhibited more exploratory behaviors and appeared to employ different response strategies. Specifically, subjects pressed the non-illuminated lever significantly more times during cued-choice trials for the lever signaling reward omission than during cued-choice trials for the rewarded lever. Rats also spent more time in the back of the operant chamber, changed behavioral focus more frequently, and spent a greater amount of time rearing during periods of reward omission. Together, these behavioral results support the idea that uncertain environments should bias behavior towards exploration and alternative strategies.

This exploration-exploitation trade-off has led to a theory in which DA is proposed to regulate the balance between exploration and exploitation (Humphries et al., 2012). Specifically, Beeler and colleagues have suggested that DA modulates the target of behavioral energy along an explore-exploit axis and a conserve-expend axis. Reward is argued to determine the distribution of energy expenditure across these axes (Beeler et al., 2012). This theory suggests that when DA is high, energy is expended and behavior focused toward exploration. Conversely, when DA is low, behavioral energy is conserved and current knowledge is exploited. A stream-lined version of this hypothesis is that tonic levels of DA regulate thrift; specifically, increasing DA decreases thriftiness, whereas decreasing DA increases thriftiness (Beeler, 2012). Supporting this hypothesis, when given a choice between a lower effort lever and a high effort lever, hyperdopaminergic (DA transporter knock-down) mice will lever press the higher effort lever more than wild-type mice (Beeler et al., 2010). Importantly, learning deficits do not appear to account for differences in behavior between the two strains of mice.

If tonic levels of DA regulate behavior across the explore-exploit axis, changes in these behaviors would be expected to correspond to the changes in [DA]. Consistent with this prediction, we observed significant correlations between [DA] and three measures of exploratory behavior, specifically, the percentage of time rearing, amount of time spent in the back of the operant chamber, and number of times in which rats switched behavioral activities. However, no significant relationship between [DA] and different response strategies (such as choosing the non-illuminated lever or ceasing to respond) was observed. Overall, these results are consistent with the interpretation that DA may control the balance between exploration and exploitation (Humphries et al., 2012); however, future studies are needed to determine the causal role of manipulations of tonic DA transmission on these specific behaviors.

Conclusion

The current study tested a number of predicted relationships between DA and behavior and, importantly, changes in response to manipulations of reward availability. While we have shown that tonic levels of DA increase during reward omission, at least under the current experimental parameters, [DA] does not purely track reward rate, since [DA] remains elevated during many of the rewarded trials as well. However, [DA] in the NAc does correspond to changes in motivational vigor and a number of exploratory behaviors. Together, these findings support a role of tonic DA in the NAc in behavioral flexibility (Beeler et al., 2014).

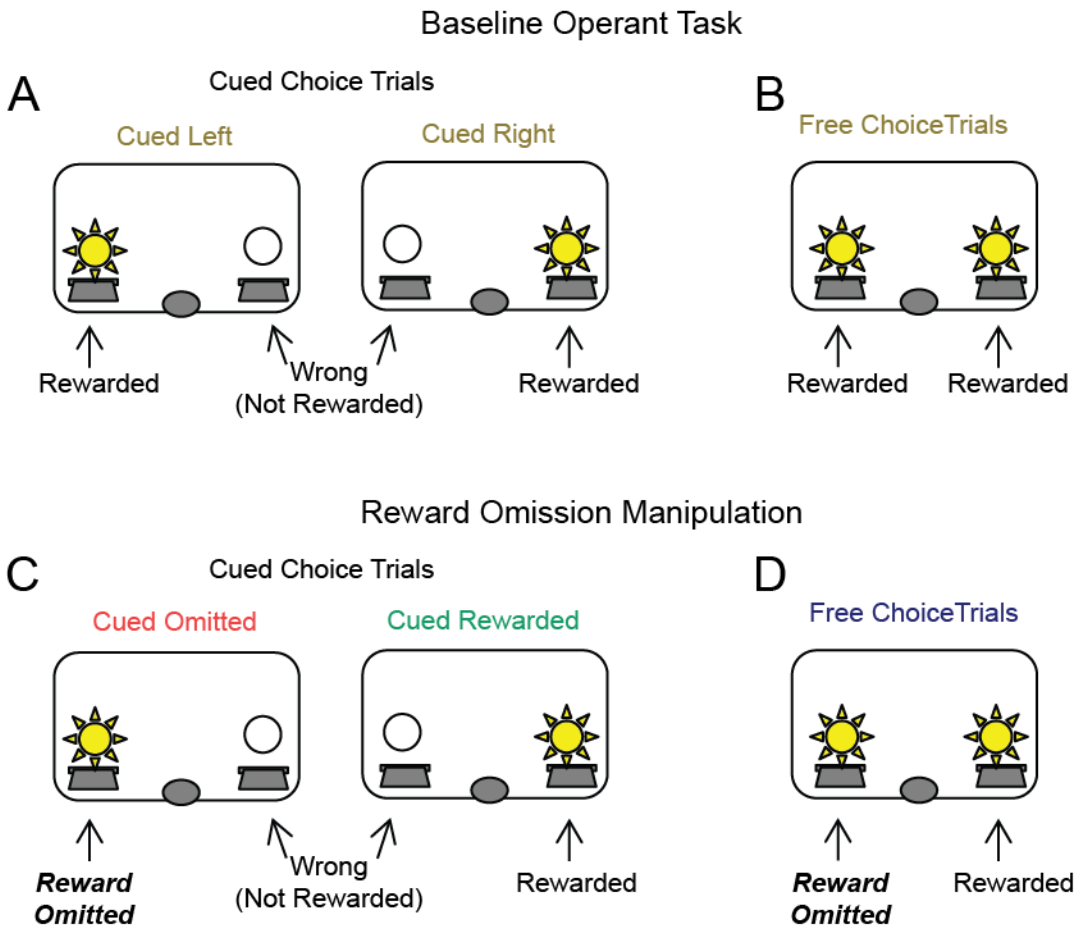

Figure 3.1. Operant, reward seeking task to model foraging behavior. **A&B**) During the baseline version of this task, rats could earn 1 reward pellet by choosing the lever under the illuminated cue light on cued choice trials (A) or by pressing either lever on free choice trials (B). **C&D**) During the reward omission session, the reward normally resulting from a correct response on one lever is omitted (C). The same lever will cease to be reinforced during the free choice trials as well (D). During the microdialysis session, all subjects received 18 trials of the baseline version of the task (A-B). Then, the reward omission manipulation began (C-D), which persisted for the duration of the session. A control group of animals received the rewarded, baseline task (A-B) for the entire session.

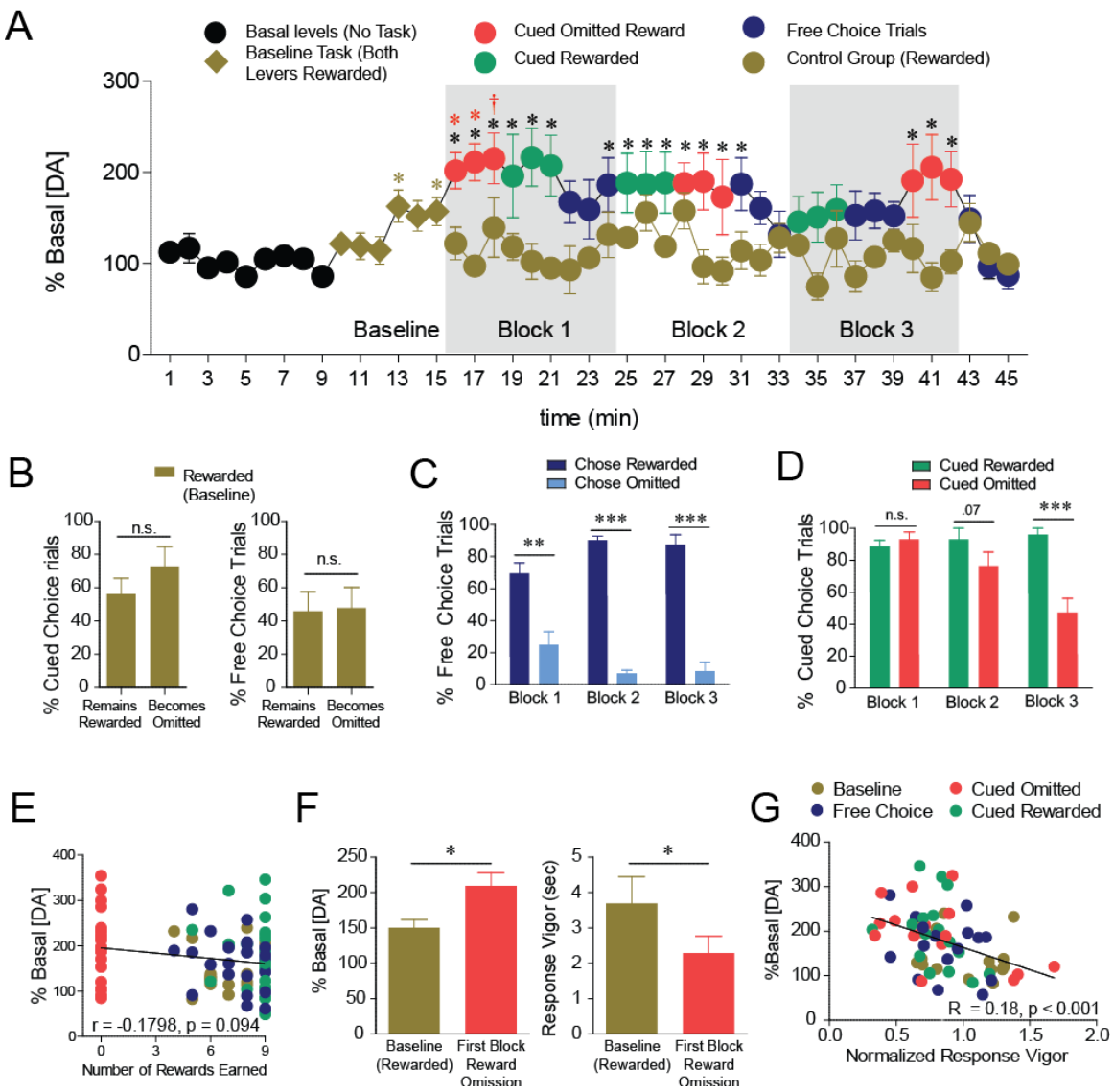

Figure 3.2. Dopamine transmission, choice behavior, and response vigor during the switch from the rewarded task to reward omission. **A**) Initially, [DA] increases above basal levels during the baseline, rewarded version of the task. Then, there is a further increase in [DA] following unexpected reward omission, which persists throughout the majority of the session. **B-D**) Choice behavior during the test session. During the baseline trials, subjects press both levers equally (B). However, following the reward omission manipulation, rats develop a robust preference for the rewarded option (C) and decrease responding on cued-choice trials for the non-rewarded option (D). **E**) The amount of reward earned does not account for changes in [DA]. **F-G**) However, the changes in [DA] do correspond to changes in vigor (defined as amount of time it takes to complete lever pressing). When [DA] increases during the first block of reward omission, response vigor simultaneously decreases (F). Furthermore, there is a correlation between

[DA] and response vigor throughout the session and across trial types (G). Error bars represent standard error of the mean. n.s. = not statistically different, *p < 0.05, **p < 0.01, ***p <0.001.

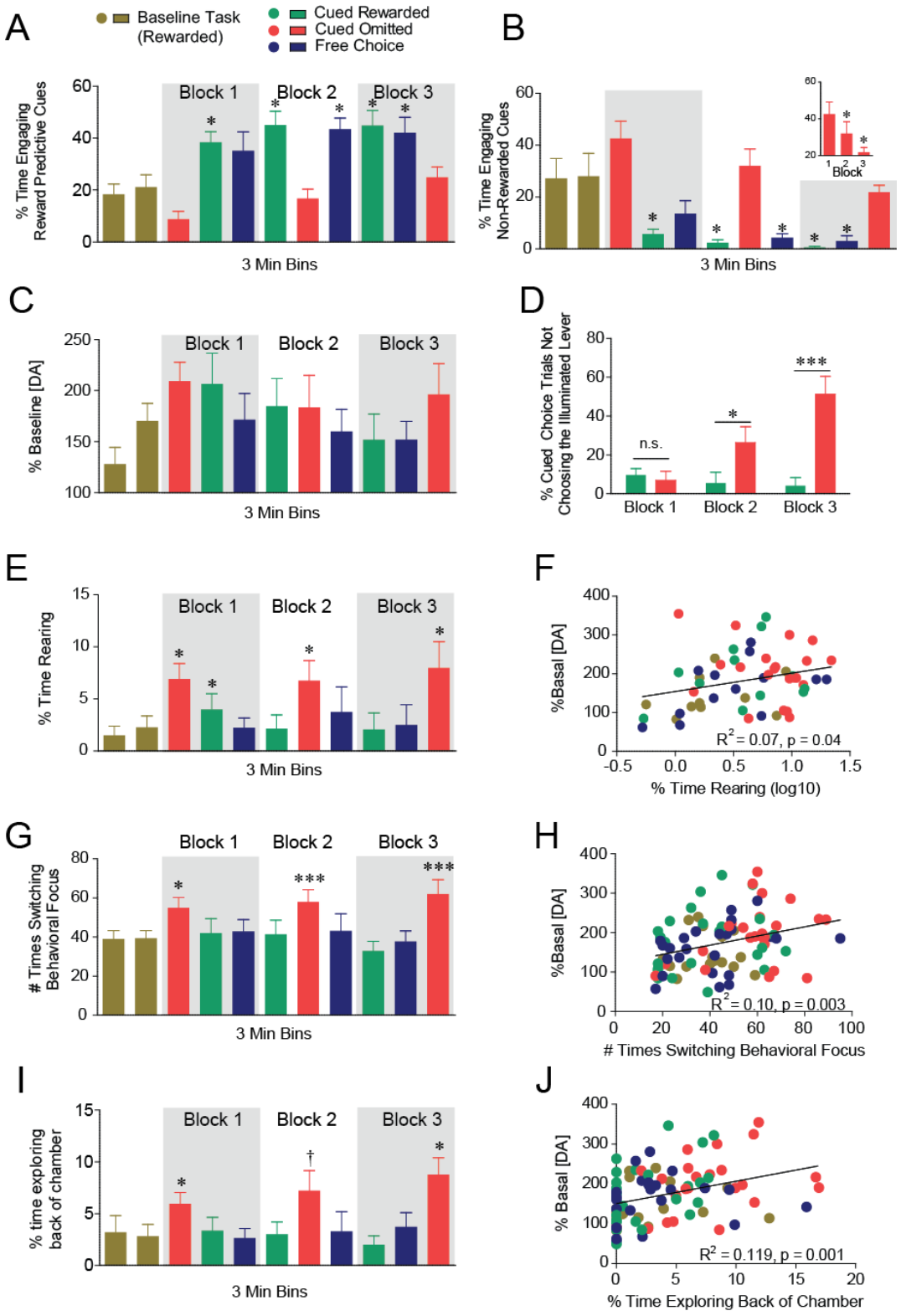

Figure 3.3. Changes in [DA] correspond to behavioral allotment between the exploitation of knowledge and the exploration of new options. **A**) Amount of time spent engaging the reward predictive cues. **B**) Percentage of time engaging the cues signaling

no reward availability. **C**) Changes in [DA] organized in 3 min bins to facilitate comparisons with behavior. **D**) Percentage of cued choice trials in which subjects made "incorrect" responses, which includes pressing the non-illuminated cue light or not pressing either lever. **E**) Percentage of time subjects exhibited rearing behavior. **F**) Relationship between rearing behavior and [DA]. **G**) The number of times subjects changed their behavioral focus from one activity to another. **H**) Correlation of switching behavioral focus with changes in [DA]. **I**) Changes in percentage of time spent exploring the back half of the chamber away from the cues and food port. **J**) Relationship between [DA] and time exploring the back half of the chamber. Error bars represent standard error of the mean. n.s. = not statistically different, † = 0.076, *p < 0.05, **p < 0.01, ***p <0.001.

CHAPTER 4:

PHASIC CHANGES IN DOPAMINE TRANSMISSION IN THE NUCLEUS ACCUMBENS CORE FOLLOWING UNEXPECTED REWARD OMISSION

Introduction

In chapter 3, we demonstrated that reward omission causes an increase in tonic levels of dopamine (DA) in the nucleus accumbens (NAc) lasting many minutes, and that these increases correspond to changes in motivational vigor and exploratory behaviors. While such increases in DA concentration ([DA]) may be necessary to invigorate the animal to search for reward elsewhere, the temporal precision of these changes is not adequate to signal which lever is being extinguished. Here, we explore the DA transmission dynamics occurring on a sub-second time course during the same behavioral paradigm.

Based upon a number of elegant electrophysiology studies recording from a sub-population of putative midbrain DA neurons, these neurons have been proposed to signal reward prediction errors (Schultz, 2002). When an unexpected reward is received or an outcome is better than expected, DA neurons have been shown to burst fire; whereas, when an outcome is worse than predicted, these neurons can exhibit a pause in firing (Schultz et al., 1997; Matsumoto and Hikosaka, 2009). These observations have provided a potential neurobiological substrate for temporal difference learning. Specifically, these phasic changes in neural activity have been proposed to mediate

learning about the change in reward (Schultz, 2002; Bromberg-Martin et al., 2010b; Steinberg et al., 2013) or the incentive value of rewards (Flagel *et al.*, 2011; Berridge, 2012). While most of the electrophysiology studies have been done in non-human primates, these prediction errors have also been observed in rats (Hyland *et al.*, 2002b; Roesch *et al.*, 2007), and similar effects have been detected in the human ventral striatum using fMRI to detect changes in BOLD signal (McClure *et al.*, 2003; Haruno *et al.*, 2004; Abler *et al.*, 2006), suggesting a common mechanism across these species.

Fast-scan cyclic voltammetry (FSCV) studies recording changes in extra-cellular [DA] have revealed that a number of DA transmission dynamics in the NAc are consistent with predictions made about DA release by electrophysiological studies. For example, when an unexpected reward is received, a phasic increase in DA release occurs, termed reward-evoked DA, and if a cue predicts the reward, as animals learn the relationship between the cue and reward, the increase in [DA] transfers to the reward-predictive cue (Day *et al.*, 2007). These patterns of DA transmission have been observed for Pavlovian conditioned stimuli (Day *et al.*, 2007; Flagel *et al.*, 2011; Clark *et al.*, 2013) as well as discriminative stimuli requiring an operant response to obtain the reward (Day *et al.*, 2010; Gan *et al.*, 2010; Wanat *et al.*, 2010; Sugam *et al.*, 2012). Very recently, DA release in the NAc has been shown to encode the reward prediction error term (Hart *et al.*, 2014); however, it is noteworthy the timing of the detected reward prediction error DA signal in that study occurred 2 sec after reward onset (later than electrophysiology studies would predict).

Furthermore, the magnitude of cue-evoked DA has been shown to track the relative utility of currently available options. For example, cue-evoked DA can be higher

for discriminative stimuli signaling the availability of a larger reward versus an option in which less reward is available (Gan *et al.*, 2010). Additionally, cue-evoked DA has been shown to track effort-related costs, whereby there is greater cue-evoked DA to discriminative stimuli signaling a low effort option versus a high effort option (Day *et al.*, 2010).

Here, using fast-scan cyclic voltammetry (FSCV), we examine changes in DA release in the NAc during our reward omission paradigm. We hypothesize that the omission of an expected reward may be coded through decreases in DA in the NAc. Additionally, we hypothesize that the cue-evoked DA for the extinguished option would decrease throughout the session if it signals the value of the available outcomes. At first, the magnitude of cue-evoked DA would be expected to be equivalent on all cued-choice trials, since the available reward from both discriminative stimuli has been equal throughout training; however, as the animal experiences and learns about the reward omission, we hypothesize that the magnitude of the cue-evoked DA on the extinguished lever would decrease, compared to cue-evoked DA on rewarded trials, throughout the session.

Methods

Subjects

Sprague Dawley rats obtained from Charles River (Winington, MA) were used as subjects for this study. Once rats were trained on the behavioral task (exactly as described in chapter 2), they underwent surgery. Out of 34 rats that underwent FSCV surgery and testing, 8 subjects rendered usable data and are included in this experiment.

Animals were kept on a 12:12 hr reverse light/dark cycle (lights off at 8am) and tested during their dark cycle. Prior to surgery, rats were pair house, and after surgery, subjects were single housed. Subjects had free access to water in their home cages and were mildly food restricted to approximately 90% of free-feeding weight (Porter-Stransky *et al.*, 2013). Subjects were fed daily after testing.

Surgery

All procedures were approved and performed in accordance with the University of Michigan Committee on the Use and Care of Animals. Subjects were anesthetized with an intramuscular injection of ketamine hydrochloride (90 mg/kg) and xylazine hydrochloride (10 mg/kg). Rats were also given the analgesic ketoprofen (5 mg/kg) and the antibiotic cefazolin (70 mg/kg).

A guide cannula (Bioanalytical Systems, West Lafayette, IN) was positioned over the NAc core (relative to bregma: 1.4 mm anterior and 1.3 mm lateral) and cemented into place. Additionally, a Ag/AgCl reference electrode (6-7 mm long) was implanted into contralateral cortex. Then, a carbon fiber microelectrode was acutely lowered through the guide cannula into the dorsal portion of the NAc (15 turns), and a bipolar stimulating electrode was lowered into the VTA (relative to Bregma: 5.2 mm posterior and 0.8 mm lateral). To optimize placement of the stimulating electrode, stimulations of the VTA (60 pulses at 60 Hz, 120 μA) were given while recording the electrically-evoked DA release in the NAc at a carbon fiber microelectrode, as previously described (Porter-Stransky *et al.*, 2011). The average optimized position of the stimulator was 8.6 mm ventral (relative to top brain surface). Stimulators were cemented in place to facilitate electrically-evoked DA release on the day of testing, which was necessary for the chemometrics used to

convert recorded current into [DA] (Keithley *et al.*, 2010). The carbon fiber recording electrode was removed after the stimulator was cemented into place, so that a fresh recording electrode could be used the day of testing.

Fast-scan cyclic voltammetry

Carbon fiber microelectrodes encased in glass capillary (exposed carbon fiber 105-125 µm) were pre-calibrating using known concentrations of DA in a flow cell (Sinkala *et al.*, 2012). On the day of testing, these recording electrodes were lowered through the guide cannulae into fresh brain tissue in the NAc. Depth into brain is monitored by voltammetrists by the number of turns the electrodes is lowered through the standard micromanipulator (Robinson and Wightman, 2007); the average site of recordings was 17.5 turns into brain (recording sites are depicted in Fig. 4.1). A triangular waveform ramping from -0.4 V to +1.3 V and back to -0.4 V was applied against the implanted Ag/AgCl reference electrode at 60 Hz for 20 min and then 10 Hz for an additional 10 min to facilitate electrode stabilization. Throughout the experiment, changes in current were recorded 10 times per second.

Since FSCV captures changes in current from the oxidation and reduction of a variety of electro-active compounds, changes in current specifically due to DA can be extracted using principle components regression (Heien *et al.*, 2004; Keithley *et al.*, 2010). To facilitate this process, electrically-evoked DA release was elicited by stimulation of the VTA before and after the experiment, as previously described (Porter-Stransky *et al.*, 2011). Additionally, a few unexpected pellets were given to subjects, since these are known to reliably evoke transient increases in [DA] in the NAc (Day *et al.*, 2007; Clark *et al.*, 2013). Details about the chemometric analyses are provided below.

Behavioral testing

Behavioral testing during FSCV recordings was conducted as described in chapter 3. Subjects first completed 18 trials of the baseline version of the task (Fig. 3.1A-B). Then, the reward omission manipulation began where one lever was extinguished while the other lever continued to be reinforced as usual (Fig. 3.1C-D). To facilitate comparison with the results of the microdialysis experiment, trials were blocked during FSCV testing the same way as during the microdialysis experiment (Fig. 3.2A).

Histology

After FSCV experiments, lesions were made at the same depth in the brain as the recording occurred (Robinson and Wightman, 2007), since carbon fiber microelectrodes do not leave tissue damage detectable with light microscopy (Khan and Michael, 2003; Peters et al., 2004). Brains were sliced 50 µm thick on a cryostat. Electrode placements were examined under 10x magnification with a light microscope (Leica, Buffalo Grove, IL). Recording sites are depicted in Figure 4.1.

Data Analysis

Conversion of recorded current into dopamine concentration

Recorded current was converted into changes in [DA] using chemometric analysis with principle components regression (Heien *et al.*, 2004; Keithley *et al.*, 2009). Since both DA and pH can both substantially contribute to changes in recorded current, DA and pH were included into the chemometric analysis. DA traces were obtained from electrically-evoked DA release *in vivo* and pellet-evoked DA (by giving an unexpected reward to the rat). Following electrically-evoked DA release, a shift in pH is often

observed (Michael *et al.*, 1998; Phillips *et al.*, 2003). This conveniently provides example traces in pH for each electrode. Examples of both acidic and basic changes in pH were included into the pH component of the chemometric analysis, since both are observed during the experiments. Malinowski's F-test was used to determine the number of factors to keep in the regression model (Keithley *et al.*, 2010).

Since FSCV works through background subtracting out a large current, traces are depicted as changes (Δ) in [DA]. Consistent with previous studies (Day *et al.*, 2010), data "snippets" were created for each trial and aligned to the illumination of the cue lights. As a rule, the lowest point in the 5 sec prior to cue light illumination was chosen as the point to background subtract. Therefore, bar graphs indicating DA values are relative to this pre-cue point. Any data points that exceeded the Qα value in the residual analyses were excluded (Keithley *et al.*, 2010). To examine changes in DA during reward delivery and reward omission, snippets of changes in [DA] for each trial were realigned to lever retraction, which occurred immediately after the operant response was completed.

Statistical Analyses

Changes in [DA] traces between reward delivery and reward omission were calculated using a linear mixed model because of its ability to properly handle correlated, repeated measures data (Aragona *et al.*, 2008; Aragona *et al.*, 2009). The baseline value for detecting significant changes in [DA] during reward delivery and reward omission was the [DA] value when the operant response concluded (time point 0 in Fig. 4.3B).

Cue-evoked DA was quantified in two ways. Consistent with previous studies (Gan *et al.*, 2010), the peak cue-evoked DA was determined by the highest point within 2

sec of the cue light illuminating (although, here, the peak was generally observed 0.3 - 0.6 sec after the onset of the cue light, as seen in Figs. 4.2-4.6). Additionally, since DA release has been shown to last up to a few seconds after a behaviorally-relevant event (Wanat *et al.*, 2010; Wanat *et al.*, 2013; Hart *et al.*, 2014), we examined changes in DA during the 5 sec period while the cue lights were illuminated before levers were available by quantifying area under the curve (AUC). For comparisons in which there were only two conditions, paired t-tests were used.

To examine changes in behavior, cue-evoked DA, and relative [DA] during reward delivery versus reward omission across the 3 blocks, repeated measures ANOVAs were used with Bonferroni corrections applied to post hoc tests. Subjects whom did not have data in one or more conditions in the time course analyses were excluded from these comparisons, as not to skew the within subject statistical tests. For all analyses, and α level of 0.05 was used, and statistical analyses were performed using IBM SPSS Statistics 21 (Armonk, NY) and GraphPad Prism (La Jolla, CA).

Results

To examine changes in sub-second DA transmission during unexpected reward omission, FSCV was conducted in the NAc of rats. Fig. 4.1 depicts the recording sites. Placements of the carbon fiber microelectrode were in the NAc core for 7 rats, and 1 placement was on the borderline of the medial NAc shell. Data from this recording site did not appear different from those fully in NAc core, so the shell placement was combined with the core placements and included in the analyses below.

During the FSCV test session, subjects first received 18 baseline trials in which both levers were equally reinforced. Subjects completed cued choice trials for both the left and right levers equally (Fig. 4.2A; $t_{(7)} = 0.189$, $p = 0.856$). Similarly, on free choice trials, subjects pressed both levers, not demonstrating a significant preference for one lever over the other (Fig. 4.2B; $t_{(7)} < 0.001$, $p = 1.000$). The average changes in [DA] are shown in Fig. 4.2C. As expected, peak cue-evoked DA release did not significantly differ across baseline trial types (Fig. 4.2D; $F_{(2,12)} = 0.615$, $p = 0.557$). Finally, there was not a statistically significant change in [DA] when the reward pellets were delivered (Fig. 4.2E; linear mixed model, $p > 0.05$ at all time points).

After the baseline trials concluded, the reward omission manipulation began. During this portion of the session, one lever was extinguished while the other lever continued to be reinforced. The same lever was reinforced on cued choice and free choice trials for each subject, although which lever was extinguished was counter-balanced across subjects. Consistent with the experiment in chapter 3, subjects reduced responding on cued choice trials for the non-rewarded lever over the course of the session (Fig. 4.3A; main effect, $F_{(1,19)} = 66.385$, $p < 0.001$). Initially, before subjects had learned about the extinguished lever, subjects completed both trial types equally (block 1, $p = 0.583$). Then, throughout the second and third blocks, subjects completed significantly fewer non-rewarded cued choice trials (Fig. 4.3A; blocks 2 and 3, $p < 0.001$).

During block 1 when subjects were experiencing unexpected reward omission for the first time, there was a significant decrease in [DA] when the anticipated reward was omitted (Fig. 4.3Bi.; linear mixed model, $p < 0.05$ for time points 3.0 – 4.6 on graph, $p < 0.1$ at time points 2.2-2.8 and 4.8). On cued rewarded trials, a brief increase in [DA] was

detected when the reward was delivered (Fig. 4.3Bi.; $p = 0.029$ at time point 2.4). During the second block, which is when subjects began to reduce responding on the non-rewarded lever, there were no significant changes in [DA] during reward delivery or reward omission (Fig. 4.3Bii.; $p > 0.05$ at all time points). By the third block, subjects rarely pressed the non-illuminated lever (Fig. 4.3A). On the rare occasion that they did choose the non-rewarded lever, a decrease in [DA] after the reward omission was seen (Fig. 4.3Biii.; $p \leq 0.05$ at time points 4.4-6.0). There were no statistically significant changes in [DA] when subjects received the rewarded in block 3 (Fig. 4.3Biii.; $p > 0.05$ at all time points). DA transmission at the moment of reward delivery on rewarded trials significantly differed from the same time point in non-rewarded trials when the reward was omitted during the first block; however, DA transmission did not statistically differ during blocks 2 and 3 (Fig. 4.3C; main effect, $F_{(1,16)} = 8.587$, $p = 0.010$; block 1, $p = 0.007$; block 2, $p = 0.453$; block 3, $p = 0.168$). Together, these data reveal phasic changes in [DA] in the NAc core that are consistent with reward prediction errors.

Fig. 4.4A shows the average changes in [DA] on rewarded and non-rewarded, cued choice trials during blocks 1 (i.), 2 (ii.), and 3 (iii.). Consistent with behavioral performance, cue-evoked DA did not initially differ between the two trial types during the first block. However, by the second and third blocks, when subjects were responding differently on these trials (Fig. 4.3A), cue-evoked DA began to differentiate between the two discriminative stimuli (Fig.4.4A-C). During blocks 2 and 3, the overall DA transmission during the 5 sec presentation of the discriminative stimulus (Fig. 4A), quantified by area under the curve, was lower to the cue light signaling no reward availability compared to the cue light signaling reward would be available (Fig. 4.4B;

main effect, $F_{(1,15)} = 6.360$, $p = 0.023$; block 1, $p = 0.985$; block 2, $p = 0.027$; block 3, $p = 0.077$). Additionally, the peak cue-evoked DA was significantly different between rewarded and non-rewarded cued choice trials during by the last block (Fig. 4C; main effect, $F_{(1,15)} = 4.323$, $p = 0.055$; block 1, $p = 0.567$; block 2, $p = 0.123$; block 3, $p = 0.022$). Together, these results demonstrate that cue-evoked DA in the NAc tracks the changing value of discriminative stimuli.

In chapter 3 we showed that during the cued, non-rewarded trials, subjects employed alternative response strategies, such as pressing the non-illuminated lever (Fig. 3.3D). Here, we replicated this effect; on some cued, non-reward trials subjects chose the illuminated lever, while other times choosing the non-illuminated lever (Fig. 4.5A). While neither response resulted in a reward, rats did choose the illuminated lever more often than the non-illuminated lever (Fig. 4.5B; $t_{(6)} = 2.804$, $p = 0.031$). DA transmission differed on these cued choice trials depending on which lever the subject choose (Fig. 4.5C). Even before the rats could press the lever, the peak cue-evoked DA was higher on trials in which subjects subsequently pressed the illuminated lever than on trials in which the pressed the non-illuminated lever (Fig. 4.5D, $t_{(6)} = 2.804$, $p = 0.031$).

During the free choice trials, subjects could choose either the rewarded lever or the non-rewarded lever (Fig. 4.6A). Consistent with our previous experiments, rats developed a strong preference for the rewarded option during free choice trials (Fig. 4.6B; $F_{(1,19)} = 89.507$, $p < 0.001$), that was evident during all three blocks (block 1, $p = 0.019$; block 2, $p < 0.001$; block 3, $p < 0.001$).

DA transmission during the free choice trials was very similar to that on cued choice trials for the rewarded lever and was higher than on cued choice trials for the non-rewarded option (Fig. 4.6C). The DA transmission on the trials was separated into two groups based upon the animals' responses (choosing the rewarded versus the non-rewarded option). Since subjects rarely chose the non-rewarded lever during blocks 2 and 3, there were not enough trials to do proper time course analyses on the free choice trials, so the data below are from all three blocks combined. [DA] appeared less on the non-rewarded trials compared to the rewarded trials (Fig. 4.6D); however, there was more variation in average [DA] in the non-rewarded trials (likely due to a much smaller sample size of trials in which they chose the non-rewarded option; Fig.4.6B). The peak cue-evoked [DA] for free choice trials did not significantly differ between trials in which subjects choose the rewarded lever versus the non-rewarded lever (Fig. 4.6E).

Discussion

Decreased DA transmission during the omission of an expected reward

Using FSCV, we examined the sub-second DA transmission dynamics in the NAc as subjects experienced the omission of an expected reward and developed a behavioral preference for the rewarded option. Consistent with the predictions make by electrophysiology studies of putative midbrain DA neurons (Schultz et al., 1997; Bromberg-Martin et al., 2010b) and a recent FSCV study (Hart *et al.*, 2014), DA transmission decreased when expected rewards were omitted (Fig. 4.3B-C).

DA has been hypothesized to function as a teaching signal, detecting when violations of expectation occur (Hollerman and Schultz, 1998). If DA detects violations

of expectation, then the decreased DA transmission during reward omission would be predicted to dissipate as subjects learned about this contingency switch and the omission was no longer surprising. Importantly, the reduced [DA] during reward omission was most evident during the first block of the session, when the omission of reward was most unexpected, and this effect disappeared by the second block (Fig. 4.3B-C). Critically, the "negative prediction error" in the DA corresponds to learning measured by changes in behavior. Specifically, the decrease in [DA] was strongest during the first block when subjects were responding and expecting reward; however, it disappeared by the second block when subjects decreased responding on the non-rewarded lever. This finding is consistent with electrophysiology experiments revealing that the decreases in DA neuronal activity to reward omission lessen as the omission becomes expected (Roesch *et al.*, 2007).

Some have questioned whether dips in [DA] in terminal regions from a pause in the firing rate of DA neurons are detectable or substantial enough to function as a reliable learning signal (Niv and Schoenbaum, 2008). Indeed, the baseline firing rate of classical DA neurons in the rat is low, less than 10 Hz (Hyland *et al.*, 2002a; Pan *et al.*, 2008), and the magnitude of increases in firing rate can be much greater than decreases in firing rate. To elucidate if the small reductions in DA transmission during the omission of an expected reward are functionally relevant for the observed changes in behavioral response to reward omission, we are conducting a follow up experiment utilizing optogenetics in transgenic rats whereby a brief optical stimulation is given during the reward omission to counteract the dip in DA transmission on non-rewarded trials.

Since the phasic decreases in [DA] are small in magnitude compared to phasic increases, the downstream mechanisms that signal this information are likely quite different. Phasic increases in DA release are thought to stimulate low-affinity, D1-like receptors, whereas phasic decreases in [DA] are hypothesized to be signaled through decreased binding at high-affinity, D2-like receptors (Bromberg-Martin et al., 2010b; Dreyer et al., 2010). Indeed, modeling data have shown that phasic decreases in [DA] briefly reduce D1- and D2-like receptor occupancy to 0% (Dreyer *et al.*, 2010). While phasic decreases in DA may affect both families of DA receptors, the decreased binding at D2-like receptors is likely the receptor class that is functionally affected, since these receptors have a high affinity to DA and high baseline occupancy (Richfield *et al.*, 1989). Supporting this mechanism of action, in chapter 5 we show that preventing a decrease in DA tone at D2-like receptors through site-specific microinfusion of D2-like, but not D1-like, agonists dose-dependently prevents the behavioral preference for the rewarded option (Porter-Stransky *et al.*, 2013).

It is important to note that while the predominant view in reinforcement learning is that phasic decreases in [DA] signal a negative prediction error that could function as a teaching signal, this is not the only plausible explanation. The change in [DA] to the omission of an expected reward could potentially signal other qualities such as motivational disappointment (Berridge, 2012), frustration (Amsel, 1958, 1962), or even aversion (Daly, 1974; McCutcheon *et al.*, 2012). Indeed, DA transmission in the NAc has been shown not to be necessary for all types of learning (Palmiter, 2007). Utilizing naturally occurring individual differences in the propensity to attribute incentive value to reward predictive stimuli, it has been shown that DA is specifically involved in the

attribution of incentive motivational qualities to cues (Flagel *et al.*, 2011). Therefore, reduced DA transmission during the reward omission could be signaling a decrease in the incentive value of the non-rewarded option and the cues signaling that option.

Much as the time course of decreased DA transmission during non-reward attenuates as the session progresses, and as it has been used to support DA's role in learning, the same logic could be applied to the alternative theories of DA. The reduction of incentive value, disappointment, or frustration could all be expected to be most intense during the first few experiences of the omitted reward (i.e. during block 1) and reduce intensity through the session, mapping onto the time course of changes in DA transmission throughout the session (Fig. 4.3B-C). Therefore, while emphasis has been biased to the learning theories of DA since they have provided clear predictions of phasic DA transmission, there are definitely other plausible, competing interpretations for the role of phasic DA transmission in the NAc. Much follow up research is needed to parse apart the specific functions of phasic DA transmission in the NAc on motivated behavior.

Changes in cue-evoked DA transmission as subjects learn about the extinguished option and alter behavioral responses

Since many of the original electrophysiology studies used simpler, non-choice tasks, there had been a paucity of research on if and how DA codes decision making and choice behavior (Niv *et al.*, 2006). Over the past few years, more studies have incorporated choice tasks to determine the relationship between DA and specific actions. Some studies have revealed that cue-evoked DA codes the best available option, regardless of which option is chosen (Roesch *et al.*, 2007; Day *et al.*, 2010). Others,

however, have suggested that DA codes which option will be chosen, thereby playing a more prominent role in decision making (Morris *et al.*, 2006).

Here, we found that during free choice trials, cue-evoked DA did not differ based upon the subsequent action of the rat (Fig. 4.6E). This finding is consistent with other studies indicating that DA signals the best available option (Roesch *et al.*, 2007; Day *et al.*, 2011; Sugam *et al.*, 2012). However, an important caveat is that subjects chose the non-rewarded lever significantly less often during the free choice (Fig. 4.6B), and, specifically in this paradigm, subjects rarely choose the non-rewarded lever during free choice trials after the first block (Fig. 3.2C). Therefore, while these data appear most consistent with the notion that cue-evoked DA signals the best available option, this finding does not preclude the possibility that cue-evoked DA could differentiate between options. Indeed, there were not enough free choice trials in which subjects divided their responding between the two options across the session blocks to perform time course analyses on free choice trials. For this very reason, cued choice trials were included in the test session. Since subjects continue to respond during cued choice trials for the non-rewarded lever (Fig. 4.3A) longer than during free choice trials (Fig. 4.6B), these cued choice trials provide a larger sample to compare changes in DA and behavior between rewarded and non-rewarded trials across the session.

During the first third of the test session, cue-evoked DA was equal between the rewarded and non-rewarded trials (Fig. 4.4B-C), during which subjects were still responding equally on rewarded and non-rewarded cued choice trials (Fig. 4.3A). As subjects gained experience with the reward omission, cue-evoked DA began to differentiate between the two options and changes in behavioral responses emerged.

Specifically, subjects reduced responding on the non-rewarded lever (Fig. 4.3A), and DA transmission decreased to the cue light signaling no available reward compared to the cue light signaling reward availability (Fig. 4.4B-C). These data demonstrate that as learning about contingency changes occurs (or as the non-rewarded option loses incentive value), the magnitude of cue-evoked DA updates, reflecting the new value of the available options.

To elucidate if the magnitude of cue-evoked DA predicted behavioral response within a trial type, we analyzed differences in cue-evoked DA on non-rewarded trials when subjects pressed the illuminated lever (which was considered a correct response and rewarded during training) versus pressing the non-illuminated lever (which was an incorrect, non-rewarded response during training and testing). Peak cue-evoked DA significantly differed when the cue-evoked DA was grouped by subjects' subsequent responses on each trial (Fig. 4.5). On average, cue-evoked DA was higher when subjects choose the illuminated lever than the non-illuminated lever (Fig. 4.5C-D). This demonstrates a relationship between variations in cue-evoked DA for a single cue and the response strategy that subjects would choose for a given trial.

Conclusion

Here, we demonstrated that sub-second changes in [DA] in the NAc code when violations of expected outcomes occur. Additionally, DA transmission to cues predicting the ability of reward updates as subjects learn about the changing value of the options signaled by these cues. Together, the changes in cue-evoked DA in this study demonstrate that as reward contingencies change, this DA signal in the NAc can differentiate among the value of the attainable options within trials based on the cues

presented in each trial. In these ways, DA transmission in the NAc appears to be both predictive of what outcomes are available (and in some cases which option will be chosen) and reactive when expected outcomes do not occur.

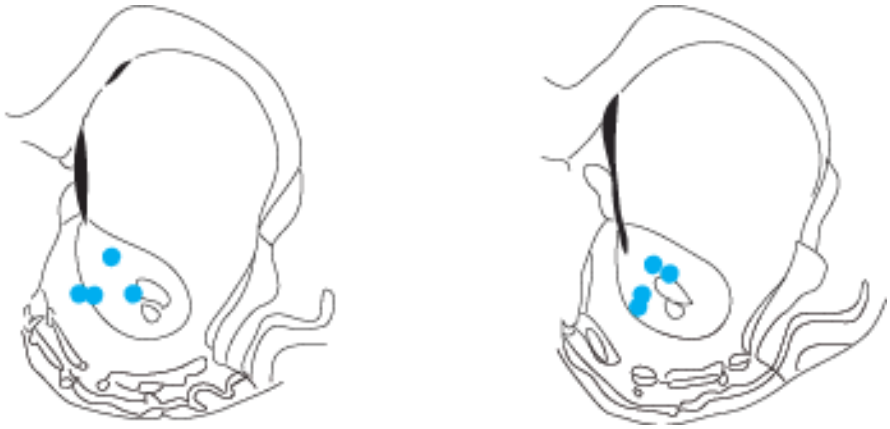

Figure 4.1. FSCV recording sites in the NAc (n = 8).

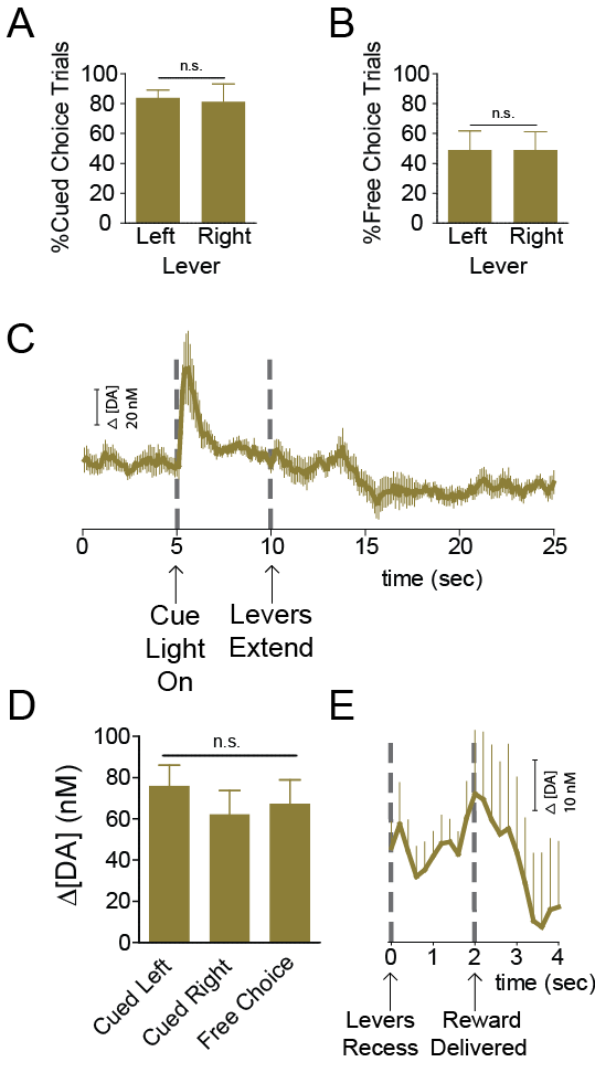

Figure 4.2. Baseline behavior and DA transmission prior to reward omission. **A&B**) Behavioral performance on cued choice trials (A) and free choice (B) trials. **C**) Average changes in [DA] during baseline trials aligned to cue light onset (at 5 sec into file). **D**) Average peak cue-evoked DA relative to before cue onset. **E**) Average changes in [DA] during reward delivery (for all trial types). Graphs indicate mean ± SEM, n.s. = not statistically different.

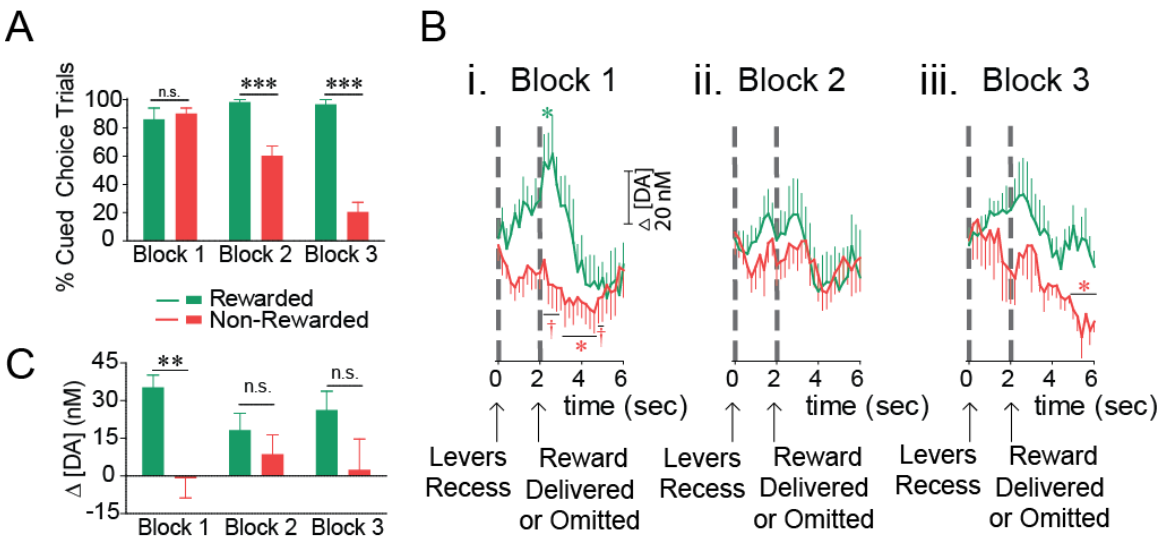

Figure 4.3. DA transmission on cued choice trials during the omission of a previously expected reward. **A**) Percentage of cued choice trials in which subjects responded on the illuminated lever. Throughout the session, subjects consistently completed cued choice trials for the rewarded lever (green) with near perfect accuracy; however, they significantly decreased responding on cued choice trials for the non-rewarded lever (red). **B**) Average DA transmission aligned to the completion of the operant response (time 0). On rewarded trials, the pellet was delivered 2 sec after the levers recessed. **C**) Changes in [DA] relative to a pre-cue baseline during the moment of reward delivery or omission. Graphs indicate mean ± SEM, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.05$, † $p < 0.1$, n.s. = not statistically different.

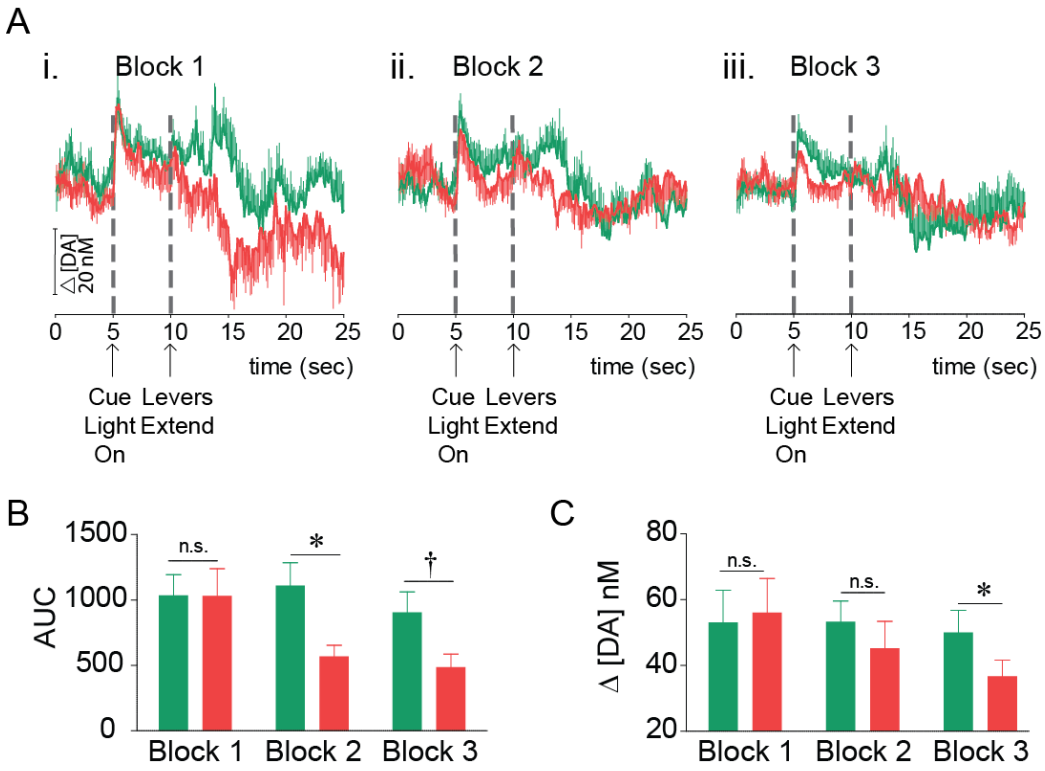

Figure 4.4. Time course of changes in cue-evoked DA transmission between rewarded and non-rewarded cued choice trials. **A**) Average DA traces during cued rewarded trials (green) and cued non-rewarded trials (red) during block 1 (i.), block 2 (ii.), and block 3 (iii.). B) Differences in cue-evoked DA measured by area under the curve during the 5 sec before the levers became available (i.e. between the dotted gray lines in A for each block). C) Peak cue-evoked DA during each block. Graphs represent mean ± SEM, n.s. = not statistically different, † $p = 0.077$, *$p < 0.05$, * $p < 0.01$, *** $p = 0.001$.

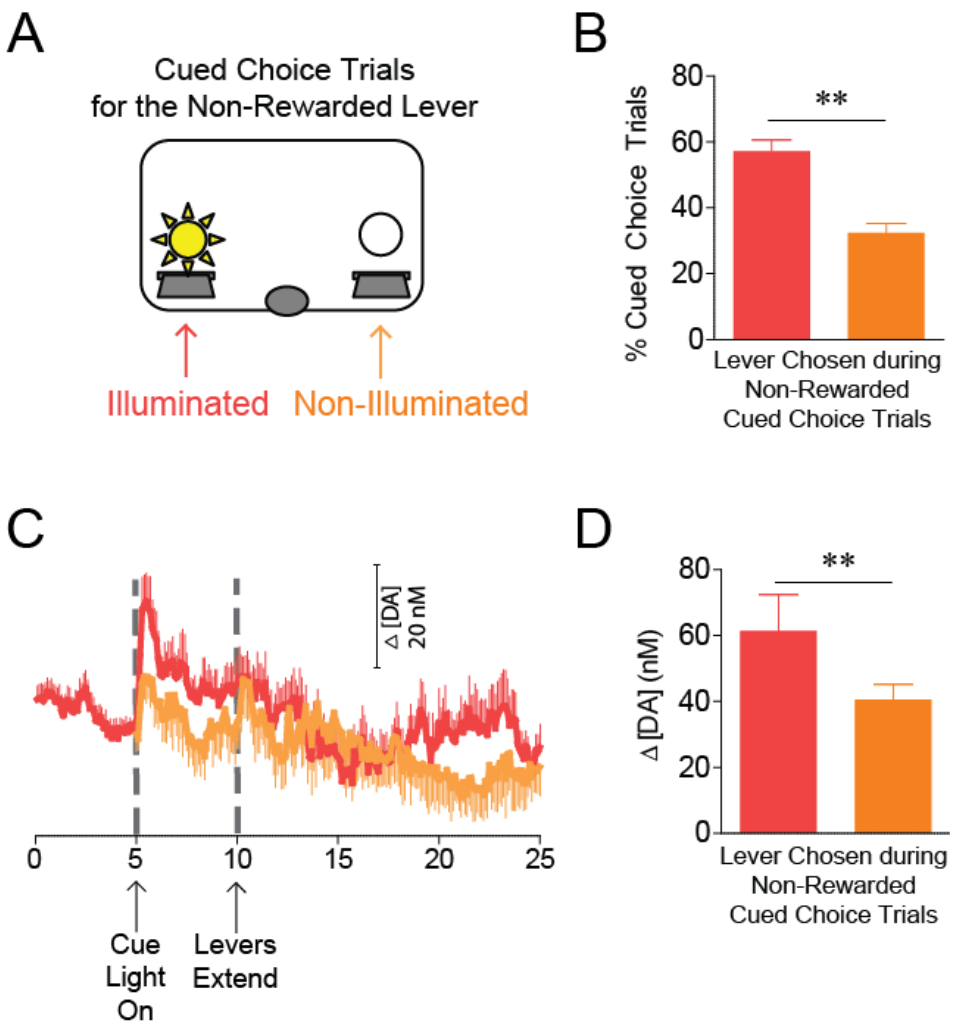

Figure 4.5. Differential DA transmission on non-rewarded, cued choice trials correspond to behavioral response. **A**) Two possible behavioral responses on non-rewarded, cued choice trials. **B**) Response allocation on the two available levers during the cued choice trials for the non-rewarded lever. **C**) DA transmission on non-rewarded, cued choice trials. **D**) Average peak cue-evoked DA on non-rewarded cued choice trials based upon subject's subsequent response on each trial. Graphs represent mean ± SEM, n.s. = not statistically different, ** $p < 0.01$.

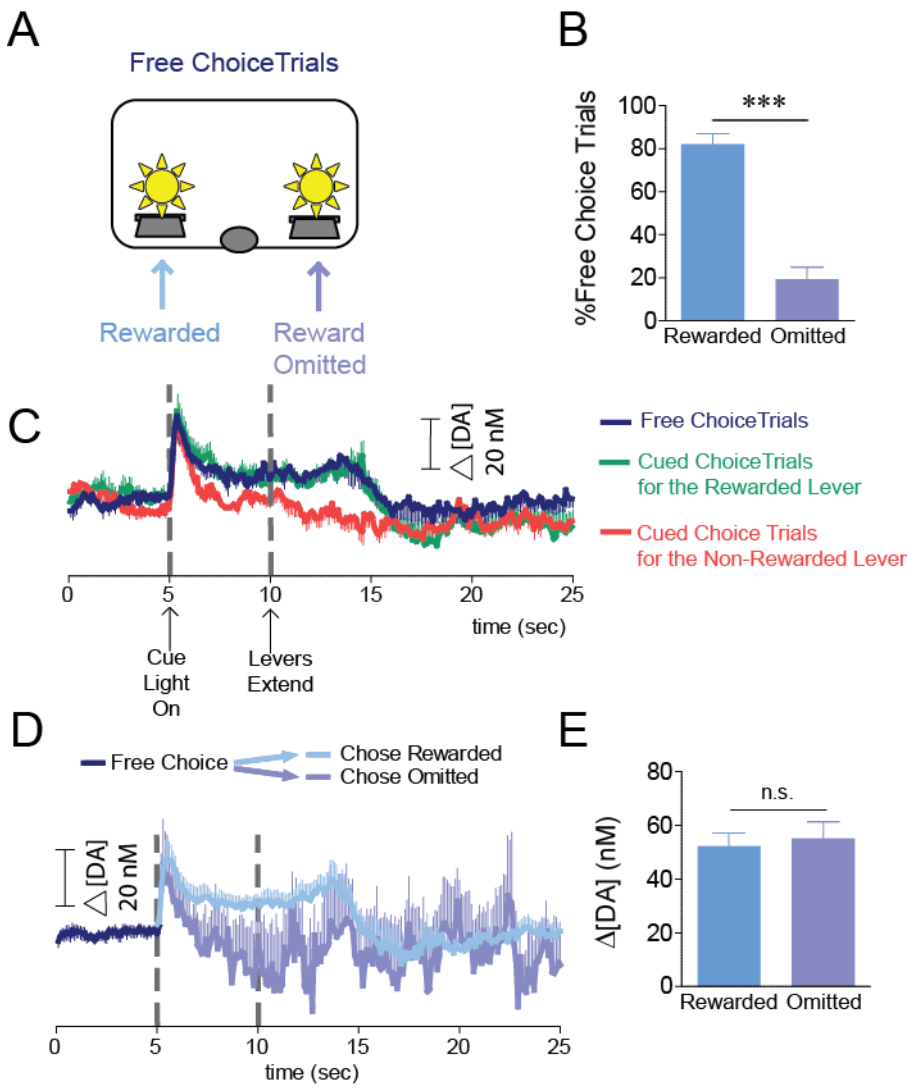

Figure 4.6. Changes in DA and behavior during free choice trials. **A**) The two available response options during free choice trials. **B**) Behavioral preferences during free choice trials. **C**) Comparison of DA transmission during free choice and rewarded and non-rewarded cued choice trials (all blocks combined). **D**) DA transmission during free choice trials based upon which lever subjects chose. **E**) Average peak cue-evoked DA on free choice trials.

CHAPTER 5:

REDUCTIONS IN D2 RECEPTOR TONE IN THE NUCLEUS ACCUMBENS MEDIATE BEHAVIORAL PREFERENCES FOR THE OPTIMAL CHOICE FOLLOWING UNEXJPECTED REWARD OMISSION

Introduction

Neurobiologically, mesolimbic dopamine (DA) has been strongly implicated in motivated behavior (Nicola, 2007; Berridge, 2012). Although DA has long been known to be involved in appetitive, reward-seeking behaviors (Schultz, 1998; Brown and Peters, 2004; Phillips et al., 2007a; Dalley and Everitt, 2009), there is growing evidence that mesolimbic DA is also involved in aversive motivation (Young, 2004; Anstrom et al., 2009; Badrinarayan et al., 2012; Salamone and Correa, 2012). Previous studies have demonstrated that the reduction or omission of an expected reward is a salient and even aversive event that can significantly alter behavior (Tinklepaugh, 1928; Miller and Stevenson, 1936; Amsel, 1958; Daly, 1974; Kerfoot et al., 2008), and aversive responses to reward omission are phylogenetically ancient (Vindas *et al.*, 2012). While the nucleus accumbens (NAc) core has been shown to mediate behavioral flexibility (Cardinal et al., 2001; Corbit et al., 2001; Floresco et al., 2006a; Haluk and Floresco, 2009), little is known about the role of DA in this system following decreased responding when reward is omitted (Annett et al., 1989; Reading and Dunnett, 1991).

Electrophysiological recordings demonstrate phasic reductions in firing rate by conventional putative DA neurons [projecting to the NAc core (Ikemoto, 2007; Lammel et al., 2008)] when an expected reward is omitted (Schultz et al., 1997; Roesch et al., 2007), and this is believed to cause a phasic decrease in DA concentration ([DA]) in terminal regions. Modeling data demonstrate that these phasic decreases reduce D1- and D2-like receptor occupancy to 0% (Dreyer et al., 2010). The impact of reducing DA receptor tone has been understudied, and which DA receptor subtype impacts behavior following unexpected reward omission remains unknown. However, phasic decreases in [DA] are hypothesized to preferentially alter D2-like receptor occupancy, since these receptors have greater affinity for DA (Richfield et al., 1989), and therefore a higher baseline occupancy (Dreyer et al., 2010; Marcellino et al., 2012). It has therefore been suggested that behavioral alterations resulting from phasic decreases in [DA] are mediated by D2, but not D1, receptors (Frank, 2005; Bromberg-Martin et al., 2010b).

Here, we utilized the same behavioral task as used in the previous three chapters allowing subjects to "forage" for reward in two different locations (two spatially distinct levers). We tested the aforementioned hypothesis by administering D1- and D2-like receptor agonists and antagonists into the NAc core prior to the first sessions of the reward omission.

Methods

Subjects

A total of 122 male, Sprague-Dawley rats (251-275 g at beginning of experiment) were used in these experiments. Rats were obtained from Charles River Laboratories

(Winington, MA, USA) were pair-housed in transparent plastic cages with metal tops. Animals were kept on a 12:12 hr reverse light-dark cycle. Experiments were run daily between 9:00 and 17:00 during the dark phase.

Mild food restriction was employed to train rats to lever-press for the food reward. Since rats naturally continue growing, daily feeding accounted for natural growth over time, which was important to maintain consistent motivation levels throughout the experiment. Subjects were food restricted to ~90% of their free feeding weight accounting for natural growth (Baker et al., 2012). Natural growth curves for free-feeding male and female rats were obtained from Charles River Laboratories (Winington, MA). After the operant session each day, rats were weighed and fed based on their weight between 15:30 – 16:30 each day during the dark cycle. Rats had free access to water in their home cages.

Surgery

All procedures were conducted in accordance with the National Institute of Health Guide for the Care and Use of Laboratory Animals and were approved by the University of Michigan Committee on the Use and Care of Animals. Subjects in the behavioral pharmacology experiment underwent surgery after initial behavioral training. They were returned to free-feeding the day prior to surgery. On the day of surgery, rats were anesthetized with an intramuscular injection of ketamine hydrochloride (90 mg/kg) and xylazine hydrochloride (10 mg/kg) and implanted with 22 gauge bilateral stainless-steel guide cannulae (Plastics One, Roanoke, VA) above the NAc core (AP = +1.4 mm and ML = +1.3 mm relative to bregma; DV = -3.0 from surface of the skull). Guide cannulae were permanently fixed in place with two stainless-steel surgical screws and dental

acrylic. Stainless-steel obturators flush with the end of the guide cannulae were inserted until the experimental day. After surgery, all subjects were given ketoprofen (5 mg/kg) for pain relief and ad libitum access to food and water until fully recovered.

Once fully recovered from surgery (as determined by the return of normal weight gain), food restriction resumed, and subjects were re-trained on the FR4 operant task for a minimum of 4 days or until stable performance (defined as a minimum of 3 consecutive days with at least 90% accuracy on each trial type) on the task was observed. Obturators were removed, cleaned, and reinserted daily to keep the cannulae unclogged.

Drugs and Microinfusion Procedure

Since the behavioral manipulations in this study previously have been shown to phasically alter putative DA neurons (Schultz *et al.*, 1997) projecting to the NAc core (Ikemoto, 2007), we extensively examined the effects of DA transmission in the NAc core on changes in behavior resulting from reward omission. We tested two doses of a variety of dopaminergic agents. The D1-like receptor agonist SKF-38393 (0.1, 1.0 µg), D1-like receptor antagonist SCH-23390 (0.1, 1.0 µg), D2-like receptor agonist quinpirole (0.1, 1.0 µg), and D2-like receptor antagonist eticlopride (0.1, 1.0 µg) were chosen, and doses were selected based on previous studies showing these compounds to be behaviorally relevant when infused into this brain region, especially at the higher dose (Wolterink et al., 1993; Ranaldi and Beninger, 1994; Swanson et al., 1997; Pezze et al., 2007; Haluk and Floresco, 2009; Moreno et al., 2013; Stopper et al., 2013). These drugs were obtained from Sigma Aldrich (St. Louis, MO) and dissolved into sterile saline (Haluk and Floresco, 2009). Drugs were mixed fresh each day of behavioral testing (i.e. the experimental day).

On the day of behavioral testing, stainless-steel injectors (28 gauge) attached to PE-20 polyethylene tubing (Plastics One, Roanoke, VA) were inserted into the secured guide cannulae and extended approximately 3.5 mm below the tip of the guide cannulae (i.e. into fresh tissue), resulting in accurate targeting of the NAc core (see Fig. 5.1). Each subject received a bilateral infusion into the NAc core (saline control, *n* = 6; 0.1 µg SCH-23390, *n* = 6; 1.0 µg SCH-23390, *n* = 6; 0.1 µg SKF-38393, *n* = 7; 1.0 µg SKF-38393, *n* = 6; 0.1 µg eticlopride, = 7; 1.0 µg eticlopride, *n* = 9; 0.1 µg quinpirole, *n* = 8; 1.0 µg quinpirole, *n* = 5). The infusion volume of 0.5 µL per side was delivered over 60 sec via a 10 µL Hamilton syringe (Reno, NV) and Harvard Apparatus pump (Holliston, MA). Injectors remained in place for an additional 60 sec following the end of the infusion to allow the drug to diffuse; then, injectors were removed, obturators were reinserted, and behavioral testing began 10 min later (Haluk and Floresco, 2009; Hanlon et al., 2010). To ensure that observed drug effects were not attributed to drug spreading outside of the NAc core, a subset of quinpirole (i.e. the drug that produced a robust behavioral effect) subjects (*n* = 7) received the effective dose of quinpirole, except in a smaller volume (1.0 µg/0.3 µL per side), infused into the core. Additionally, to determine if the quinpirole effect was unique to the NAc core subregion or more broadly to the NAc, additional rats (*n* = 6) received the effective dose of quinpirole (1.0 µg quinpirole/0.3 µL) into the medial shell. The volume of 0.3 µL was chosen since previous work has successfully utilized this volume to study nucleus accumbens core versus shell differences in the rat (Pulvirenti et al., 1994; Pierce and Kalivas, 1995; Floresco et al., 2006a; Floresco et al., 2008). This is important, since many studies have revealed differences in core versus

shell regulation in motivated behavior (Di Chiara, 2002; Meredith et al., 2008; Reynolds and Berridge, 2008; Aragona et al., 2009).

In total, sixty-seven subjects with accurate bilateral injector placements in the NAc core were included in the analyses (saline controls, n = 6; 0.1 µg SKF-38393, n = 7; 1.0 µg SKF-38393, n = 6; 0.1 µg quinpirole, n = 8; 1.0 µg quinpirole, n = 12; 0.1 µg SCH-23390, n = 6; 1.0 µg SCH-23390, n = 6; 0.1 µg eticlopride, n = 7; 1.0 µg eticlopride, n = 9). Six subjects receiving 1.0 µg quinpirole had placements in the NAc medial shell and were included in analyses.

Locomotor Testing

Since drugs acting on DA receptors in the NAc can alter general locomotor activity, which could affect behavioral performance and therefore impact the results, we tested the locomotor effects of the higher dose (1.0 µg) of each chosen D1-like and D2-like receptor agonists and antagonists. A separate drug-naive group of rats was used so that drug infusions were made into fresh, undamaged tissue, since previous work has shown decreased spread of drug effect from repeated microinfusions (Mahler et al., 2007; Richard and Berridge, 2011). These subjects (total n = 48) were implanted with guide cannulae as described above. A between-subjects design, whereby each subject only received one drug, was utilized to exclude the possibility of sensitization effects (Henry et al., 1998; Vezina, 2004). Once fully recovered from surgery, subjects were maintained at approximately 90% of their free-feeding weight so they would be in the same motivational state as those tested in the behavioral pharmacology experiments.

Consistent with previous work (Badiani et al., 1995; Crombag et al., 1999), locomotor testing was conducted in plastic rectangular cages (45 x 24 x 18 cm) with a block in the center so rats could only explore the perimeter of the cage. These cages were equipped with photobeams to quantify two measures of locomotor activity: total number of photobeam breaks and number of crossovers, defined as moving from one end of the cage to the other. Crossovers captured locomotion across the cage and not the repetitive disruption of a single photobeam (Robinson and Camp, 1987; Paulson et al., 1991). Subjects were run in waves of 6-8 rats with saline control animals in every wave to account for any potential variation across days testing sessions (which is why more saline control rats were tested than drug treatments). Forty-nine rats received infusions, which were conducted as described above (saline controls, $n = 20$; 1.0 µg SKF-38393, $n = 7$; 1.0 µg quinpirole, $n = 9$; 1.0 µg SCH-23390, $n = 7$; 1.0 µg eticlopride, $n = 5$) and locomotor activity was monitored for 1 hr, the same length of time as the operant reward-seeking sessions. One outlier was excluded, so 48 subjects were included in analysis.

Histology

Upon completion of operant and locomotor testing, all subjects were euthanized with an overdose of ketamine delivered intraperitoneally, and brains were extracted for histological verification. After soaking in formalin solution, brains were rapidly frozen and sliced on a cryostat in 50 µm sections. Brain sections were stained with cresyl violet and viewed under 10x magnification. Placements were identified by where the end of the tract from the injector tip was located and compared to the Paxinos & Watson brain atlas (Paxinos and Watson, 1998).

Statistics

Statistical analyses were done using SPSS Statistics 19 (IBM, Armonk, NY), and data were graphed using GraphPad Prism version 5.0 (San Diego, CA). Statistical significance for all statistical tests was defined with an α level of 0.05. Bonferroni corrections were applied to post-hoc tests to reduce the risk of Type I errors (Sarter and Fritschy, 2008).

Two-way (multivariate) ANOVAs and post-hoc tests with Bonferroni corrections were used to examine the effects of drugs and doses on choice preference during free choice and forced choice trials. Since the 0.3 µL and 0.5 µL volumes of 1.0 µg quinpirole in the NAc core did not statistically differ from each other, they were combined to increase power to confidently interpret the null result. Because of the robust effect of reward omission and the increased sample size of quinpirole subjects from combining both infusion volumes in the core, the quinpirole-induced blockade of the development of a behavioral preference is interpretable, and the likelihood of it being a Type 2 error is very low. Specifically, the effect of reward omission on behavioral preference is very robust causing a statistically significant choice preference in groups with as few as five subjects, and throughout the study the robust behavioral effects of reward omission were replicated in behavior-only subjects as well as in many of the drug treatment groups.

Additionally, one-way ANOVAs were used to analyze performance on specific trial types (free choice trials choosing non-rewarded lever and rewarded forced choice trials) among drug and control conditions. Dunnett's post hoc tests were used to compare

drug groups to controls. Locomotor data were analyzed using a one-way ANOVA with planned contrasts (Gonzalez, 2009).

Results

Pharmacologically holding D2-like, but not D1-like, receptor tone in the NAc core prevents a behavioral preference for the rewarded option during unexpected reward omission

Reward omission has been shown to alter the firing of putative DA neurons (Schultz, 1998; Waelti et al., 2001; Matsumoto and Hikosaka, 2009) which should affect DA transmission in the NAc core (Ikemoto, 2007). To test whether the effects of altering DA receptor tone are indeed differentially mediated in a receptor specific manner within the NAc core, multiple doses of D1-like and D2-like receptor agonists and antagonist were microinfused into the NAc core (see Fig. 5.1A for a representative image and corresponding cartoon representation) 10 min prior to the first session of reward omission. Specifically, subjects received bilateral microinfusions of saline (Fig. 5.1B), the D1-like agonist SKF-38393 (Fig. 5.1C), the D2-like agonist quinpirole (Fig. 5.1D), the D1-like antagonist SCH-23390 (Fig. 5.1E), or the D2-like antagonist eticlopride (Fig. 5.1F). Two doses of each drug were tested (see methods for justifications of chosen doses).

If a decrease in occupancy of either D1- or D2-like receptors is necessary for establishing a preference for the better option (Dreyer *et al.*, 2010), then both the D1- and D2-like receptor agonists should block the choice preference for the more optimally rewarded option. However, if only D2-like receptors are necessary for suppressing responding to the omitted reward lever [as would be predicted by (Frank et al., 2004;

Bromberg-Martin et al., 2010b; Hikida et al., 2010)], then the D2-, but not D1-like, receptor agonist would be expected to prevent a choice preference for the rewarded lever. Multiple research groups have suggested that decreases in [DA] preferentially affect D2-like receptors, because D2-like receptors have a higher affinity for DA (Richfield *et al.*, 1989) and higher basal occupancy (Dreyer *et al.*, 2010) than D1-like receptors.

Indeed, drug treatment had a significant effect on preference behavior during free choice trials (main effect of drug treatment, $F_{(8,58)} = 3.084$ $p = 0.006$; interaction of choice preference and drug dose received, $F_{(8,56)} = 2.665$, $p = 0.015$). Comparable to behavior only subjects (Fig. 2.2B), control rats that received infusions of saline into the NAc core exhibited a robust preference for the rewarded lever on free choice trials (Fig. 5.2A; $p < 0.001$). Neither dose of the D1-like receptor agonist prevented subjects from exhibiting a significant preference for the rewarded lever (Fig. 5.2B; 0.1 µg SKF-38393, $p = 0.003$; 1.0 µg SKF-38393, $p < 0.001$). However, administration of the D2-like receptor agonist dose-dependently prevented a behavioral preference for the rewarded lever during the first session of reward omission, with the higher dose being the effective dose (Fig. 5.2C). Subjects receiving the lower dose of quinpirole developed a moderate preference for the rewarded lever ($p = 0.048$), but the higher dose of quinpirole attenuated the development of a choice preference during unexpected reward omission ($p = 0.213$). These data support the hypothesis that phasic decreases in DA transmission from unexpected reward omission (Schultz, 1998; Pan *et al.*, 2008) have functional consequences at high affinity D2-like receptors (Richfield et al., 1989) necessary for altering behavior (Frank, 2005).

In contrast to infusions of DA agonists which can functionally hold DA tone stable at specific DA receptor subtypes, administration of DA antagonists were largely without effect in this behavioral paradigm. Similarly to controls, subjects receiving either dose of the D1-like receptor antagonist displayed a significant behavioral preference for the rewarded lever (Fig. 5.2D; 0.1 µg SCH-23390, $p = 0.002$; 1.0 µg SCH-23390, $p < 0.001$). Subjects that received either dose of the D2-like receptor antagonist also developed a preference for the rewarded lever (Fig. 5.2E; 0.1 µg eticlopride, $p = 0.016$; 1.0 µg eticlopride, $p = 0.041$).

In addition to examining the effects of D1-like and D2-like agonists and antagonists on the development of behavioral preferences for the rewarded lever, we also compared how frequently subjects receiving each drug chose the omitted-reward lever. Rats receiving either dose of the D2-like, but not D1-like, receptor agonist chose the lever yielding no reward significantly more times than controls during free choice trials (Fig. 5.2F; main effect of drug treatment, $F_{(8,58)} = 2.782$, $p = 0.011$; Dunnett's post-hoc comparisons: 0.1 µg quinpirole, $p = 0.047$; 1.0 µg quinpirole, $p = 0.004$; 0.1 µg SKF-38393, $p = 0.311$; 1.0 µg SKF-38393, $p = 0.994$). To determine whether the quinpirole effect on choice preference persisted throughout the session, we divided the free choice trials into two blocks and analyzed the percentage of trials in which subjects receiving 1.0 µg quinpirole chose the non-reinforced lever compared to controls. Quinpirole subjects chose the non-rewarded lever significantly more times than controls during both blocks of trials (Fig. 5.2F inset; block 1, $t_{(16)} = 2.879$, $p = 0.011$; block 2, $t_{(16)} = 2.620$, $p = 0.019$), demonstrating that quinpirole attenuated responding for the optimal choice throughout the session.

Although no significant changes were seen in subjects receiving the D1-like antagonist (0.1 µg SCH-23390, $p = 0.514$; 1.0 µg SCH-23390, $p = 0.926$), subjects receiving the higher ($p = 0.042$) but not lower ($p = 0.162$) dose of the D2-like antagonist chose the lever yielding no reward a modest but significant number of times more than controls. However, this was likely attributable to the significantly altered locomotor activity and/or motivation resulting from eticlopride (see locomotor results below). Indeed, subjects in the high eticlopride dose group completed fewer rewarded forced choice trials than controls (Fig. 5.2G; $p = 0.047$), while no statistical differences were seen among any of the other groups ($F_{(8,58)} = 2.848$, $p = 0.010$; 0.1 µg eticlopride, $p = 1.000$; 0.1 µg SCH-23390, $p = 0.988$; 1.0 µg SCH-23390, $p = 0.702$; 0.1 µg SKF-38398, $p = 1.000$, 1.0 µg SKF-38398, $p = 1.000$; 0.1 µg quinpirole, $p = 0.919$, 1.0 µg quinpirole, $p = 1.000$). This supports the interpretation that differences in performance of eticlopride rats (from D2-like receptor blockage within the NAc core) were at least partially attributable to other effects of the drug. It is important to emphasize that we utilized such high doses, despite their locomotor impairments, to demonstrate that such manipulations have no impact on altering behavioral choice in this paradigm.

Similarly to behavior-only controls (Table 2.1), saline controls (Fig. 5.2H, $p = 0.009$) and D1-like receptor agonist groups (Fig. 5.2I; 0.1 µg SKF-38393, $p = 0.001$, 1.0 µg SKF-38393, $p = 0.015$) completed significantly fewer forced choice trials for the omitted reward lever than for the rewarded lever. The D2-like receptor agonist, however, dose-dependently altered performance on forced choice trials (Fig. 5.2J). Subjects receiving the lower dose of quinpirole completed fewer forced choice reward-omitted trials (Fig. 5.2J; $p = 0.019$), while the higher dose caused no statistically significant

reduction in performance (Fig. 5.2J; p = 0.217) on forced choice trials even though subjects were not receiving a reward on these trials. Similarly to controls, both D1-like (Fig. 5.2K, 0.1 µg SCH-23390, p = 0.015, 1.0 µg SCH-23390, p < 0.001) and D2-like (Fig. 5.2L; 0.1 µg eticlopride marginally significant, p = 0.063; 1.0 µg eticlopride, p = 0.001) receptor antagonist groups completed fewer forced choice trials for the omitted reward lever than for the rewarded lever. These results further support the interpretation that preventing a reduction of D2-, but not D1-like, receptor tone precludes the alteration in reward-seeking strategy in this task when a "foraging patch" is depleted.

Since reward omission can cause subjects to avoid the quadrant of the chamber containing the omitted-reward lever (Fig. 2.6A), videos of the sessions of subjects receiving the higher dose of each drug were analyzed to determine the effects of DA receptor agonists and antagonists on which quadrants of the chamber rats were occupying throughout the reward omission sessions (see Fig. 2.1A). Similarly to behavior-only controls (Fig. 2.6A), saline controls spent significantly more time in the quadrant containing the rewarded lever than the quadrant containing the non-reinforced lever (Fig. 5.2M; p = 0.011). As expected, subjects receiving the D1-like agonist also spent significantly less time in the omitted-reward lever quadrant than the rewarded lever quadrant (Fig. 5.2N; p = 0.013). However, subjects receiving the D2-like agonist spent equivalent time in the quadrants containing the rewarded and omitted levers (Fig. 5.2O; p = 0.333). Also, similarly to controls, subjects infused with the D1-like (Fig. 5.2P; p = 0.004) or D2-like (Fig. 5.2Q; p = 0.012) antagonists spent significantly less time the quadrant containing the non-reinforced lever compared to time spend in the quadrant containing the rewarded lever.

Together, these results strongly suggest that holding D2-, but not D1-like, receptor tone stable prevents a choice preference for the reinforced option following a reward omission manipulation, and these data are consistent with previous work revealing that higher (1.0 or 10.0 µg) but not lower (0.1 µg) doses of quinpirole into the NAc core impairs strategy set shifting and reversal learning, which are other important assays of behavioral flexibility (Haluk and Floresco, 2009). Although basal, steady-state levels of DA have been estimated to be in the low (6-20) nanomolar range (Kawagoe et al., 1992; Sam and Justice, 1996; Shou et al., 2006; Owesson-White et al., 2012), the higher dose of quinpirole is likely better maintaining D2-like receptor tone to prevent the putative phasic reduction in D2-like receptor occupancy during reward omission (Hong and Hikosaka, 2011). Our findings support the hypothesis that a phasic reduction in D2-like receptor signaling is necessary for guiding motivated behavior away from suboptimal choices (Frank, 2005; Bromberg-Martin et al., 2010b; Dreyer et al., 2010).

To investigate whether the quinpirole-induced lack of a preference for the rewarded option during reward omission (Fig. 5.3A) was due to impairment in learning about the omitted reward, rats were tested drug-free in a second reward omission session the following day. If quinpirole caused a learning deficit, quinpirole subjects when tested drug-free would be expected to choose the suboptimal option more times than controls since they would have to learn about the contingency switch during the second session. Conversely, if quinpirole impaired the expression of the behavioral preference without impairing learning, quinpirole subjects when tested drug free would be expected to immediately perform as well as controls during the second, drug-free session of reward

omission, since learning would have occurred even though the expression of the behavioral preference was not observed during the first session (Fig. 5.3A).

During the second reward omission session when subjects did not receive any microinfusions, a robust behavioral preference was observed for the rewarded lever during free choice trials (Fig. 5.3B, main effect, $F_{(1,59)} = 1{,}032.841$, $p < 0.001$; Bonferroni post hoc results of all drug conditions was $p < 0.001$). Regardless of drug treatment the previous day, subjects chose the omitted reward lever infrequently on free choice trials (Fig. 5.3C); no significant differences in percentage of free choice trials choosing the omitted reward lever were observed compared to controls ($F_{(8,60)} = 1.146$, $p = 0.347$). Additionally, subjects did not differ in number of times they chose the omitted reward option during the first five free choice trials of the second session compared to controls (Fig. 5.3D; $F_{(8,60)} = 1.629$, $p = 0.136$), demonstrating that all subjects, including those who had received quinpirole, performed similarly during the beginning of the second session regardless of previous drug treatment. These results support the interpretation that the D2-like agonist attenuated the behavioral preference for the optimal choice during reward omission without impairing learning about the omitted reward.

Both the NAc core and shell subregions facilitate the development of a behavioral preference during unexpected reward omission through reduction of DA tone at D2-like receptors

Subregions of the NAc, primarily the NAc core and shell, are anatomically and functionally distinct (Kelley, 1999; Zahm, 1999; Di Chiara, 2002; Aragona et al., 2006; Ikemoto, 2007; Aragona et al., 2008; Aragona, 2011). The NAc core and shell subregions can serve different roles in certain types of behavioral flexibility (Floresco et

al., 2006a). Since the NAc shell can facilitate memory of arousing experiences, such as a significant reduction in expected reward (Kerfoot et al., 2008), and it recently has been shown that under certain experimental conditions aversive stimuli phasically decrease [DA] in the shell (McCutcheon *et al.*, 2012), we investigated whether the effect described above, demonstrating that the D2-like receptor agonist prevents the development of a behavioral preference for the rewarded option, is unique to the NAc core or whether it is also true in the NAc shell.

Since 1.0 µg quinpirole (0.5 µL volume) in the NAc core blocks the development of a choice preference during reward omission (Fig. 5.2C), we infused this effective dose, but at a smaller volume (0.3 µL), into the NAc core or medial shell prior to the first session of reward omission (Fig. 5.4A; core, $n = 7$, shell, $n = 6$). Neither the main effect ($F_{(1,11)} = 2.818$, $p = 0.121$) nor the interaction of choice behavior by subregion ($F_{(1,11)} = 0.152$, $p = 0.704$) was significant, demonstrating that D2-like agonism in both the NAc core and medial shell attenuates a choice preference for the rewarded option following unexpected reward omission (Fig. 5.4B). Indeed, while control subjects displayed a robust preference for the rewarded lever ($t_{(7)} = 5.347$, $p = 0.001$), subjects receiving quinpirole into the core ($p = 0.363$) or shell ($p = 0.186$) did not significantly prefer the rewarded lever compared to the non-reinforced lever.

General locomotor-effects of quinpirole are not responsible for the lack of choice preference following unexpected reward omission

To examine if the quinpirole-induced blockade of the development of the behavioral preference during reward omission was attributable to alterations in locomotor activity caused the drug, separate groups of drug-naive rats received bilateral

microinfusions of saline or the D1-like or D2-like receptor agonists or antagonists (1.0 µg/0.5µL) into the NAc core. Each subject received only one drug to exclude the possibility of sensitization effects (Henry et al., 1998; Vezina, 2004). Locomotor behavior was monitored for 1 hr – the identical length of time as the operant task. Two measures of locomotor activity were recorded: the total number of photobeam breaks and the number of "crossovers" (see methods for details).

One-way ANOVAs revealed that drug treatment significantly affected locomotor activity assessed by total number of beam breaks (Fig. 5.5A; $F_{(4,42)} = 7.88$, $p < 0.001$) as well as total number of crossovers (Fig. 5.5B; $F_{(4,42)} = 6.866$, $p < 0.001$). Specifically, D1-like activation within the NAc core significantly increased (beam breaks, $p = 0.002$; crossovers, $p = 0.001$) while D1-like blockade (beam breaks, $p = 0.007$; crossovers, $p = 0.043$) and D2-like blockade (beam breaks, $p = 0.025$; crossovers, $p = 0.044$) significantly decreased the number of photobeam breaks and crossovers compared to controls (Fig. 5.5A&B). D2-like activation (via quinpirole), however, did not significantly alter locomotor activity (beam breaks, $p = 0.715$; crossovers, $p = 0.687$).

Both number of beam breaks and crossovers reveal similar effects of the drugs on locomotor activity. Indeed, there was a significant and robust correlation between these two measures (Fig. 5.5C; $r_{(45)} = 0.971$, $p < 0.001$). The similarity between beam break and crossover data within each drug treatment demonstrates that these variables are highly reliable measures of general locomotor activity and are sensitive to alterations from dopaminergic drugs.

These locomotor data demonstrate that all doses of the dopaminergic agents used were sufficient to modulate behavior in a predicted way. Consistent with the literature, both D1- (McGregor and Roberts, 1993; Baldo et al., 2002; Haluk and Floresco, 2009) and D2-like antagonists (Boye et al., 2001; Haluk and Floresco, 2009) significantly decreased locomotor activity. Conversely, the D1-like agonist significantly increased locomotor activity, consistent with previous work (Dreher and Jackson, 1989; Phillips et al., 1995; David et al., 2004). Quinpirole, as expected (Haluk and Floresco, 2009; Stopper et al., 2013), did not significantly alter locomotion. While quinpirole did not affect general locomotion here, it is noteworthy that its effects appear to be more variable: some have detected that quinpirole site-specifically infused into the NAc modestly increased (Dreher and Jackson, 1989; Phillips et al., 1995) or decreased (David et al., 2004) locomotor activity.

Since quinpirole did not significantly alter locomotor activity, the quinpirole-induced attenuation of the behavioral preference during reward omission cannot be attributed to drug effects on locomotor activity. And, importantly, the lack of effects from all other dopaminergic drugs tested in the behavioral choice paradigm was not due to insufficient doses, because the doses utilized were sufficient to alter general locomotor behavior in expected ways.

Although the dopaminergic drugs were given at behaviorally relevant doses (Fig. 5.5A&B), response latencies of subjects during free choice trials in the reward omission experiment revealed that the locomotor effects of these drugs did not affect the ability of subjects to perform the operant response (Fig. 5.5D). Only the high dose (1.0 µg) of SCH-23390 revealed a trend for increasing response latency (p = 0.062). Response

latencies of subjects receiving the other drugs and doses did not significantly differ from controls (one-way ANOVA: $F_{(8,59)} = 3.000$, $p = 0.007$; 0.1 µg SKF-38398, $p = 0.995$; 1.0 µg SKF-38398, $p = 0.970$; 0.1 µg quinpirole, $p = 0.641$; 1.0 µg quinpirole, $p = 0.994$; 0.1 µg SCH-23390, $p = 0.668$; 0.1 µg eticlopride, $p = 1.000$; 1.0 µg eticlopride, $p = 1.000$). Together, these data demonstrate that the dopaminergic manipulations on choice behavior were not attributable to general locomotor effects of the drugs.

Discussion

Reduction of D2-like receptor tone in the NAc mediates behavioral preferences for optimal choices

When an expected reward is omitted, conventional DA neurons [known to project to the NAc core (Ikemoto, 2007; Lammel et al., 2008; Liss and Roeper, 2008; Lammel et al., 2011)] phasically decrease their firing rate (Schultz et al., 1997; Schultz, 1998; Pan et al., 2005; Roesch et al., 2007; Pan et al., 2008). This is associated with a decrease in [DA] in the forebrain terminal regions to which they project (Ikemoto, 2007; Dreyer et al., 2010), such as the NAc core (Day et al., 2007). Decreases in striatal [DA] are hypothesized to have greater functional consequences to D2-like receptors because these receptors have greater affinity for DA (Richfield et al., 1989; Marcellino et al., 2012) and therefore a higher baseline occupancy compared to low affinity D1-like receptors (Frank, 2005; Bromberg-Martin et al., 2010b; Hong and Hikosaka, 2011).

A reduction in D2 receptor tone has been hypothesized to promote action suppression and No-Go learning (Frank, 2005; Bromberg-Martin et al., 2010b). D2 expressing neurons are predominantly in the indirect pathway (Gerfen and Surmeier,

2011) and have been shown to mediate aversive learning (Hikida et al., 2010). Recently, it has been shown that freezing behavior to an aversive stimulus is strongly associated with phasic decreases in [DA] within the NAc core (Badrinarayan *et al.*, 2012; Oleson *et al.*, 2012), and optogenetic depolarization of neurons expressing D2 receptors within the dorsomedial striatum [a region which shows similar DA transmission dynamics as the NAc core (Brown et al., 2011)] causes mice to instantaneously freeze (Kravitz et al., 2010). Moreover, mice will avoid a trigger activating neurons that express D2 receptors within this region (Kravitz et al., 2012). These recent studies suggest that phasic reductions in [DA] may activate D2 expressing neurons in the NAc core and may, at least in part, mediate aversive motivated behavior.

Here, using site-specific microinfusions of D1- and D2-like receptor agonists and antagonists, we found that only the D2-like agonist quinpirole dose-dependently prevented the rapid expression of a behavioral preference for the rewarded option. Importantly, the locomotor effects of quinpirole are not responsible for the lack of a behavioral preference. Administration of the D1-like agonist or D1- or D2-like antagonists did not impair the development of a preference for the rewarded lever, which is consistent with previous work demonstrating that DA blockade does not affect the ability of rats to choose a larger reward (Salamone *et al.*, 1994). These findings further support the hypothesis that guiding motivated behavior away from aversive cues is mediated through a phasic reduction in the occupancy of D2-like receptors.

Parkinson's disease, which occurs when DA neurons degenerate, can cause cognitive impairments in addition to the well-known motor impairments (Rana *et al.*, 2013). Our finding that D2-like agonists attenuate the avoidance of a sub-optimal choice

is consistent with and has implications for research examining reinforcement learning in Parkinson's patients. Specifically, when off their medications in a DA depleted state, Parkinson's patients have been observed to be better at learning to avoid negative outcomes than positive outcomes; however, when on their medications, primarily D2-like agonists, they did not learn as well from negative feedback (Frank et al., 2004; Cools et al., 2006; Cools et al., 2007; Voon et al., 2010). Importantly, fMRI studies reveal that DA agonists which disrupt learning from negative feedback in Parkinson's patients correspond to smaller decreases in ventral striatal activity in response to losses (Voon et al., 2010). In combination with our results, these data strongly support the idea that performance in avoiding suboptimal choices is attenuated by D2-like agonists filling in the phasic dips in DA from reward omission (Frank et al., 2004).

In addition to binding to post-synaptic D2-like receptors, quinpirole also binds to pre-synaptic D2 autoreceptors. Binding of D2-like agonists to autoreceptors can decrease basal levels of DA (Kalivas and Duffy, 1991; Pierce et al., 1995; Koeltzow et al., 2003) and decrease stimulated phasic DA release in the dorsal striatum (Joseph et al., 2002; O'Connor and Lowry, 2012) as well as in the NAc core and shell (Maina and Mathews, 2010). In the present study, even though basal levels of DA may be altered due to quinpirole's effects at autoreceptors, D2-like tone at post-synaptic receptors would be maintained. Indeed, the effective dose of quinpirole in attenuating the preference for the rewarded option is in the range of doses that are presumed to bind to post-synaptic D2-like receptors (Swanson *et al.*, 1997; Boschen *et al.*, 2011). Therefore, regardless of the amount of DA being released, the post-synaptic D2-like receptors would be expected not to functionally experience the reduced binding from decreased levels of DA since the

quinpirole would be bound to these receptors, and the results of this study support this interpretation.

Indeed, basal levels of DA are sufficient to inhibit D2-expressing indirect pathway neurons (Surmeier et al., 2011), and quinpirole inhibits such neurons (Hooper *et al.*, 1997). Together, in combination with the quinpirole attenuation of behavioral response to reward omission, these data are consistent with the hypothesis that decreases in DA, which would preferentially affect high-affinity D2-like receptors (Kreitzer and Berke, 2011), are necessary for altering behavior away from a sub-optimal choice.

Conclusion

The present experiments demonstrate that the omission of an expected reward is a salient, aversive event prompting a robust preference for the rewarded option. The expression of this behavioral choice preference is dose-dependently attenuated by a D2-, but not D1-like, agonist in the NAc. These results support the hypothesis that phasic reductions in [DA], as would occur during the omission of an expected reward, preferentially affect D2-like receptors. Specifically, decreased occupancy of D2-like receptors in the NAc facilitates motivated behavior that drives animals away from a non-rewarded option.

Figures

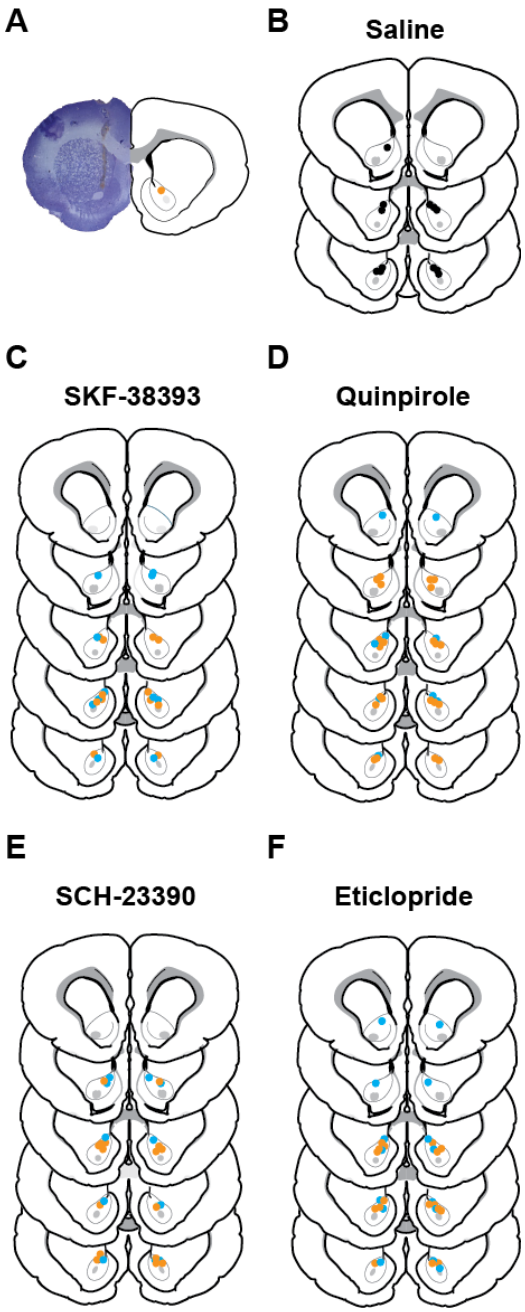

Figure 5.1. Histological verification of injectors targeting the NAc core. **A)** Representative image of injector placement and corresponding cartoon image. **B-F)** Placements of injector tips in the NAc core where drug was infused prior to the first session of reward omission. Black circles represent control saline infusions (B). Color circles indicate where the D1-like agonist (C), D2-like agonist (D), D1-like antagonist (E), and D2-like antagonist (F) were infused into the NAc core. Orange circles indicate 1.0 µg of drug, and cyan circles represent 0.1 µg of drug.

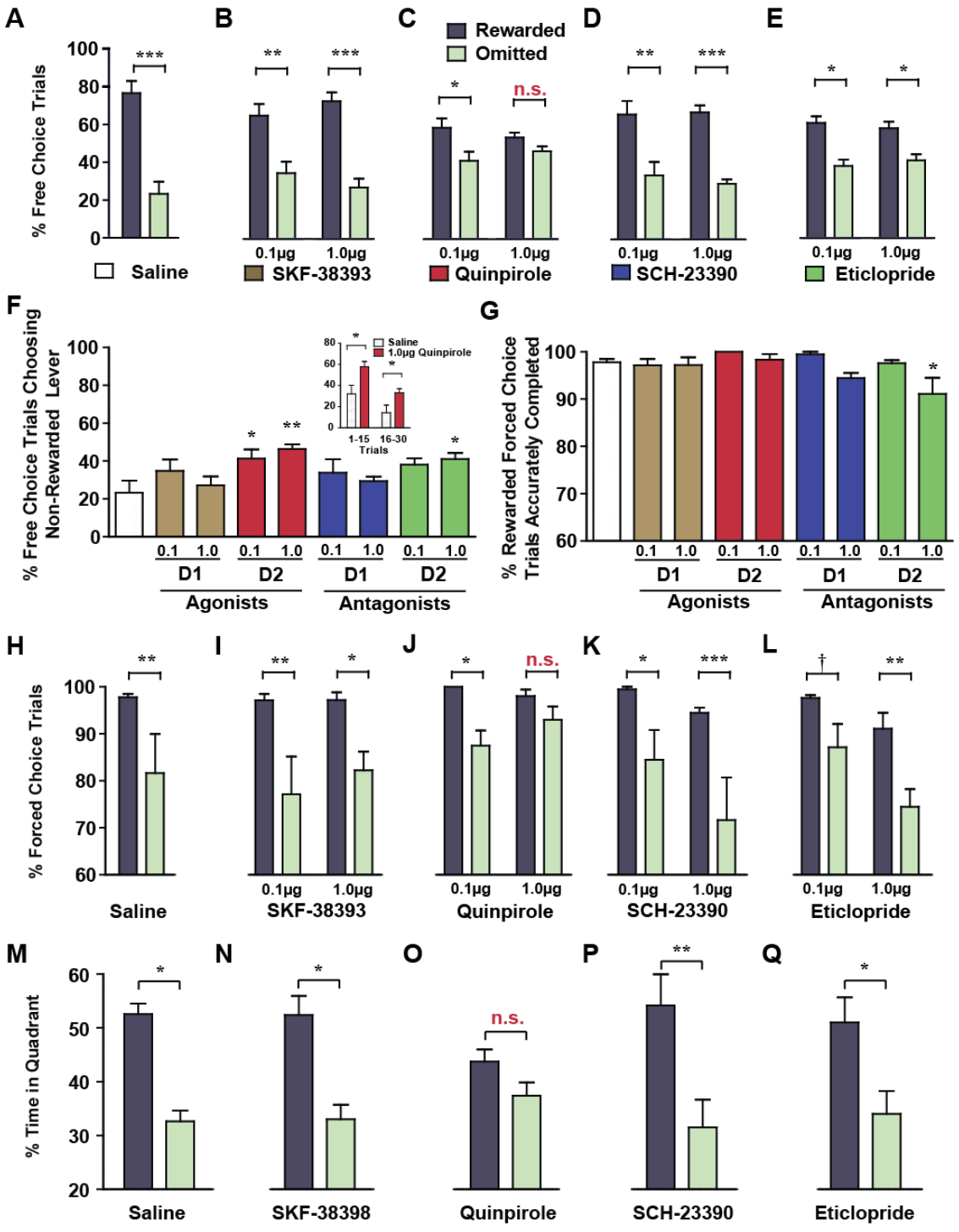

Figure 5.2. Disrupting D2-, but not D1-like, receptor tone in the NAc core prevents the rapid development of behavioral preference for the rewarded option during unexpected reward omission. **A-E)** Similarly to controls (A), subjects receiving either dose of the D1 agonist (B) or the D1 or D2 antagonists (D and E, respectively) expressed a significant preference for the rewarded lever. However, the D2 agonist dose-dependently blocked

the normally robust behavioral preference for the rewarded lever (C). **F)** Interfering with D2-, but not D1-like, receptor tone resulted in rats choosing the non-reinforced lever more often than controls (white bar) during free choice trials Inset demonstrates that the disruption of D2-like receptor tone via quinpirole affects choice preference throughout the session. **G)** Only the higher dose of eticlopride reduced the number of correctly completed forced choice trials for the rewarded lever compared to controls. **H-L)** During reward omission, subjects completed fewer forced choice trials for the non-reinforced lever than for the rewarded lever. Much like controls (H), subjects administered with either dose of the D1-like receptor agonist (I) displayed reduced performance on non-rewarded forced choice trials. The higher, but not lower, dose of the D2 agonist prevented the relative decrease in performance on non-reinforced compared to rewarded forced choice trials (J). Subjects receiving the D1-like (K) or D2-like (L) receptor antagonists completed fewer forced choice trials for the non-rewarded lever, similarly to controls. **M-Q)** During reward omission, rats avoided the quadrant of the chamber containing the omitted-reward lever, spending more time in the quadrant containing the rewarded lever. Similarly to controls (M), subjects receiving the D1 agonist (N) or either the D1-like (P) or D2-like (Q) antagonist spent less time in front of the non-reinforced lever (green bars). However, the D2-like agonist (O) prevented subjects from spending significantly more time in front of the rewarded lever (gray bars). n.s. = not statistically significant, † p=0.06, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Error bars indicate mean ± SEM.

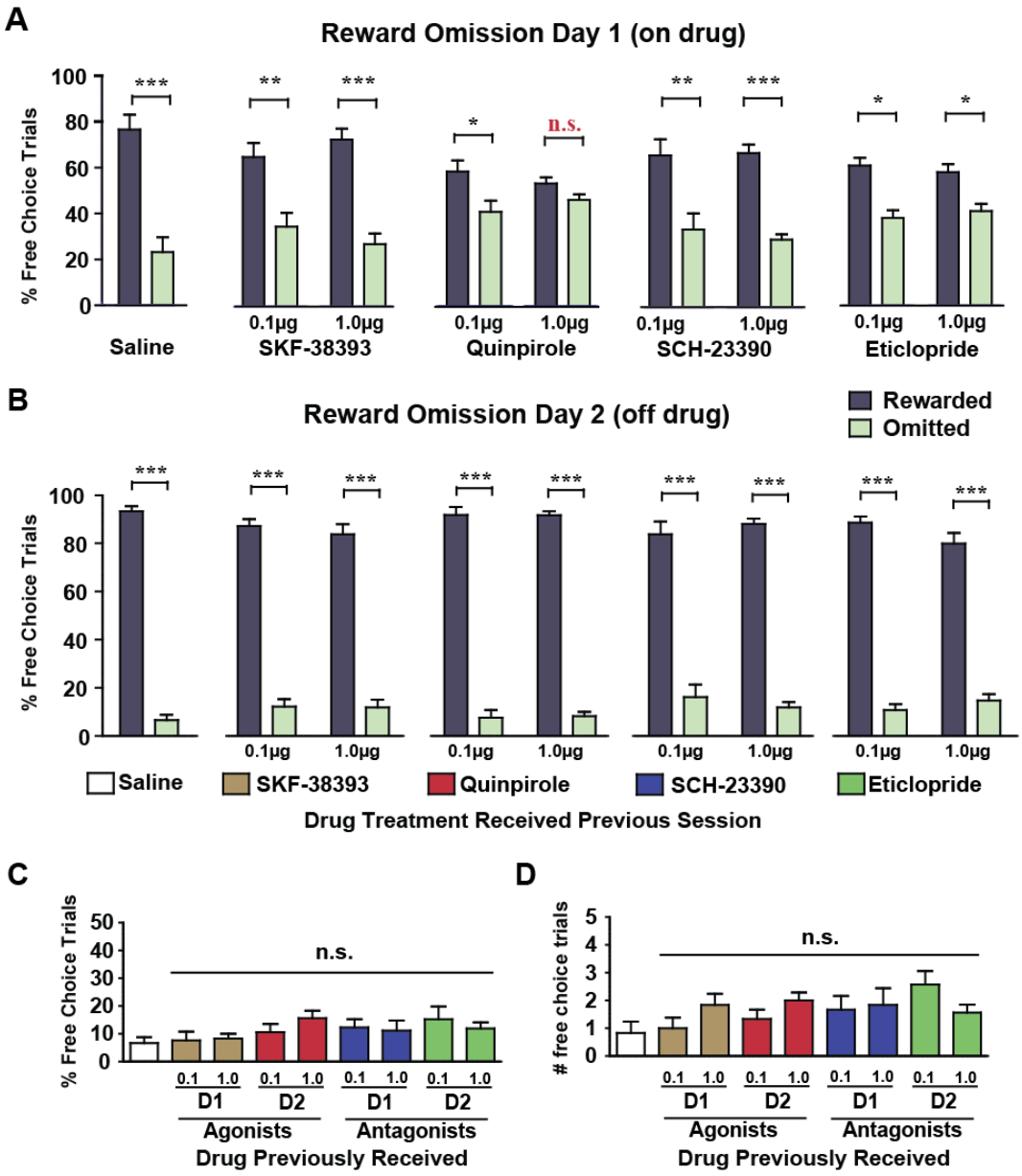

Figure 5.3. Behavioral preferences for the optimal option during the second, drug-free session of reward omission reveal that quinpirole did not block learning about the omitted reward. **A)** Choice preference from the first session is re-shown for comparison. **B)** During the second session of reward omission, rats (now drug-free) demonstrated a robust behavioral preference for the rewarded lever during free choice trials regardless of the drug they received the previous day. **C-D)** Drug received during the previous session did not significantly affect the percentage of free choice trials in which subjects chose the lever yielding no reward (C) or the number of times rats chose the omitted reward lever during the first 5 trials of the second switch day when subjects did not receive drug (D). n.s. = not statistically significant compared to saline controls, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Error bars indicate mean ± SEM.

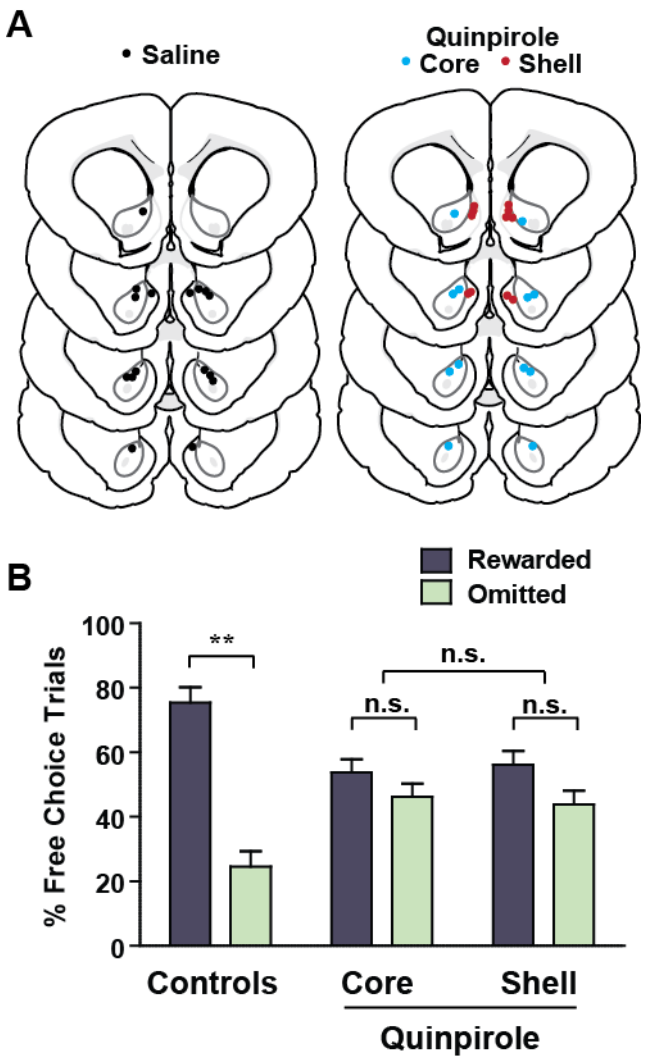

Figure 5.4. The D2-like receptor agonist administered into the NAc core or shell prevented a behavioral choice preference during reward omission. **A)** Injector placements in the NAc. Black circles represent saline controls. Red circles indicate shell and blue circles indicate core placements. **B)** 1 µg quinpirole bilaterally infused into the NAc core or shell attenuated the development of a preference for the optimal choice. n.s. = not statistically significant. Error bars indicate mean ± SEM.

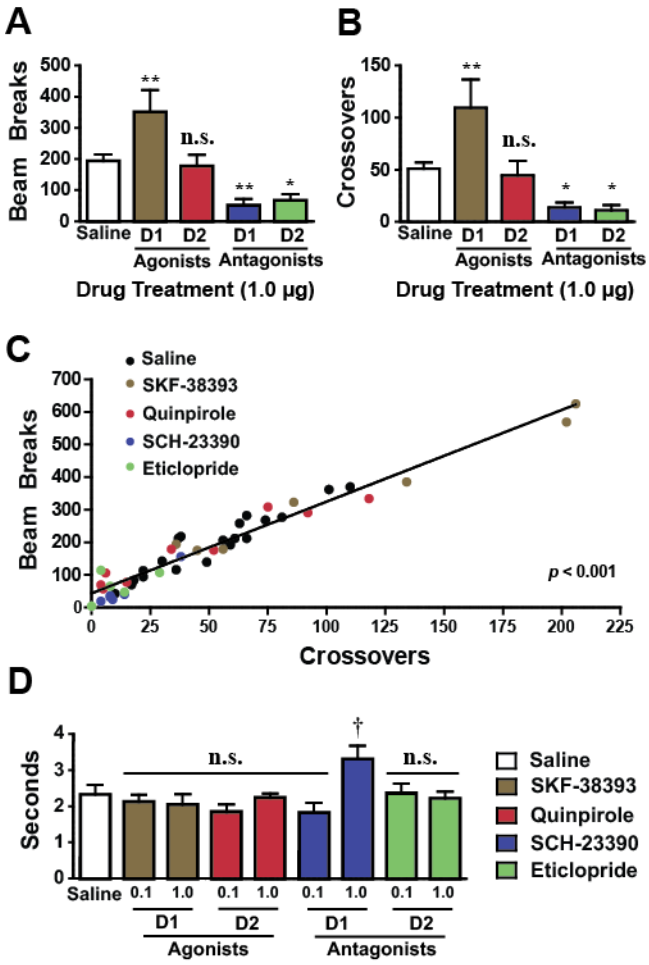

Figure 5.5. Locomotor effects of the utilized D1- and D2-like receptor agonists and antagonists. **A-B)** The D1-like receptor agonist increased while the D1- and D2-like receptor antagonists decreased locomotor activity compared to controls in separate, drug naive groups of rats as measured by total beam breaks (A) and crossovers (B) in a 1 hr session. The D2-like agonist did not affect locomotor behavior. **C)** The two measures of locomotor activity, beam breaks and crossovers, were highly correlated. **D)** Response latencies on free choice trials during reward omission reveal that only the higher dose of the D1-like antagonist affected performance on the task compared to controls (white bar). † p=0.06, *$p < 0.05$, **$p < 0.01$. Error bars indicate mean ± SEM.

CHAPTER 6:

GENERAL DISCUSSION

Foraging Behavior during Reductions in Reward Availability

In dynamic environments where resource availability frequently varies, animals must be able to identify changes in resource availability and employ adaptive foraging strategies. To model this behavior in the laboratory, in chapter 2, we modified an operant behavioral task (Day et al., 2010; Day et al., 2011; Sugam et al., 2013) to examine choice preferences in rats. On cued choice trials (also termed forced choice trials in the literature), subjects could earn a food reward by choosing the lever under the illuminated cue light. On free choice trials, the cue lights above both levers illuminated, and subjects could earn a reward by responding on either lever.

Once trained on this task, rats learned to discriminate between the cue lights with near perfect accuracy. Specifically, rats consistently learned to correctly complete the cue-choice trials, ensuring subjects were discriminating between the cue lights and responding on both levers. On free choice trials, when both levers were equally reinforced, we showed that rats sample equivalently from both levers (Porter-Stransky et al., 2013). On the test day, to model a rapid depletion, the reward following a correct response on one lever was omitted, while the other lever continued to be reinforced. During this first reward omission session, rats exhibited a strong behavioral preference for the rewarded option during free choice trials. Additionally, they decreased

responding on cued choice trials for the non-rewarded lever, while responding correctly on cued choice trials for the rewarded lever. This behavioral response to the reward omission manipulation remained consistent in both male and female rats and across all stages of the estrous cycle (Porter-Stransky et al., 2013).

In contrast to reward omission, some other negative manipulations of reward did not cause the same behavioral responses. If the reward following a correct response on one lever was reduced by 50% (instead of by 100%), rats did not show a choice preference for the optimal choice until the second reward reduction session (Porter-Stransky et al., 2013). These results are consistent with studies from other labs showing that the reduction of an expected reward causes a preference for the better option during the second session (Salinas et al., 1993; Salinas and White, 1998; Sastre and Reilly, 2006; Ramot and Akirav, 2012).

From an ecological perspective, if an animal is hungry and the food in one foraging patch has run out, it is imperative that the animal quickly adjust its foraging strategy, since successful foraging is paramount for fitness (Pyke, 1984). If, however, there still is food, albeit less food, there may not be the same sense of urgency in adapting foraging strategy. Optimal foraging theory would, however, predict that animals should adjust to the change in reward availability (Pyke, 1984), and, indeed, they do by the second session (Porter-Stransky et al., 2013).

Adaptive foraging behavior has been necessary for the survival of most species (Stephens and Krebs, 1986). In evolutionary history, foraging has been foremost critical in the direct acquisition of resources, such as food. However, the skills required for

adaptive foraging have likely paved the way for goal-directed cognition (Hills, 2006; Stout, 2010). Indeed, the processes that facilitate semantic memory (Hills et al., 2012), visual search (Wolfe, 2013), and human foraging (Hills et al., 2013) have ties to basic foraging behavior that predates the divergence between vertebrates and invertebrates (Hills, 2011).

Dopamine Transmission in the Nucleus Accumbens Following Unexpected Reward Omission

Being able to exhibit flexible, reward-seeking behavior is advantageous in the numerous aforementioned situations. Since the neurotransmitter dopamine (DA) has been heavily implicated in reward-seeking behaviors (Berridge and Robinson, 1998; Schultz, 2002; Wise, 2004; Berridge, 2012; Salamone and Correa, 2012), we conducted a series of experiments investigating DA transmission in NAc during decreased reward availability. After establishing the behavioral paradigm, as detailed in chapter 2, we sought to examine DA's involvement in the development of behavioral preferences for the rewarded option during the extinction of one foraging option. We utilized rapid-sampling microdialysis to examine changes in levels of DA over minutes in chapter 3, fast-scan cyclic voltammetry to investigate sub-second changes in phasic DA signaling in chapter 4, and site-specific microinfusions of DA receptor agonists and antagonists in chapter 5 to elucidate which DA receptor subtype mediates the behavioral preference for the optimal choice.

Changes in tonic [DA] captured by microdialysis

Using one-minute sampling microdialysis coupled with high performance liquid chromatography-tandem mass spectrometry, we examined changes in extra-cellular

levels of DA in the nucleus accumbens (NAc) throughout the instrumental foraging task. During the baseline version of the task, DA concentration ([DA]) increased to ~150% of basal levels. Then, when the reward omission manipulation began, [DA] increased to ~200% basal levels. Interestingly, [DA] remained elevated during many of the rewarded as well as non-rewarded trials of the first two blocks. However, near the end of the second block, DA transmission reduced to comparable levels to the control group (Fig. 3.2A). During the third block, [DA] significantly increased only during the three minutes of non-rewarded trials, suggesting that tonic levels of DA can adjust rapidly to environmental events (Fig. 3.2A).

Previously, it had been suggested that DA may signal the hedonic value of rewards (Wise, 2004; Wise, 2008). This seemed plausible, because [DA] has been shown to increase to a number of outcomes that animals like, such as eating, mating, and taking drugs (Berridge and Robinson, 1998). Here, although [DA] increased during the microdialysis experiment, this increase in [DA] did not correlate to the number of rewards received (Fig. 3.2E), and [DA] did not track the hedonic value of the outcomes. Indeed, rats did not appear to 'like' when an anticipated reward was omitted; in fact, rats will lever press and jump hurdles to escape stimuli associated with reward omission (Adelman and Maatsch, 1956; Daly, 1969c, b, a, 1974) and exhibit a number of behaviors indicating that reward omission can be frustrating (Amsel, 1958, 1962; Daly, 1974) and even aversive (Papini and Dudley, 1997). Additionally, manipulations of DA transmission do not alter hedonic values measured by taste reactivity (Wyvell and Berridge, 2000; Smith et al., 2011). Together, these results refute the idea that tonic levels of [DA] in the NAc track the hedonic value of outcomes.

Additionally, [DA] in the NAc over minutes did not appear to signal reward prediction errors. While phasic decreases in [DA] are predicted to occur during the specific moment that an expected reward is omitted (Schultz, 1998), tonic levels of [DA] in the NAc actually increased, rather than decreased, during reward omission (Fig. 3.2A&F). The fact that decreases in [DA] during reward omission were not detected with microdialysis is not surprising for two reasons. First, even if basal levels of [DA] remained constant except for a few brief decreases in [DA], microdialysis would likely not be able to distinguish such a short-lasting event from noise (Schultz, 2007). Second, microdialysis is likely capturing a distinct DA signal that is different from the phasic changes (Grace, 1991; Niv et al., 2005; Niv, 2007; Niv et al., 2007).

Similarly, [DA] did not track the average rate of rewards (Niv et al., 2007) or rate of punishment (Daw et al., 2002); indeed, there was an initial increase in [DA] when subjects could work for rewards during the baseline version of the task, and then there was a further increase in [DA] when one lever was extinguished and subjects could not earn rewards (Fig. 3.2).

Consistent with theories of DA's role in the vigor and effort of responding (Niv et al., 2005; Niv, 2007; Niv et al., 2007; Salamone et al., 2007), the increases in [DA] detected by microdialysis correlated to changes in the vigor of behavioral responding. Additionally, consistent with theories of DA in the balancing of behavioral exploration versus exploitation (Beeler et al., 2010; Beeler, 2012; Beeler et al., 2012; Humphries et al., 2012), the increases in [DA] during the reward omission session correlated with more time engaged in exploratory behaviors, measured by time rearing, time spent exploring the back half of the chamber away from the cues and reward port, and number of

transitions from one behavioral focus to another. While the vigor and exploration/exploitation hypotheses are theoretically framed differently, these two measures may be often be related. During periods of reward omission, invigorated behavior could be adaptive to facilitate the exploration of the changing environment and motivate the animal to find alternative avenues of acquiring food.

These neurochemical and behavioral results are also consistent with an incentive motivation interpretation (Berridge and Robinson, 1998). Pharmacologically increasing DA transmission in the NAc by site-specifically administering amphetamine increases motivated behavior toward a reward predictive cue (Wyvell and Berridge, 2000) and increases how hard animals were willing to work for food (Zhang et al., 2003). The exploratory behaviors could alternatively be interpreted as increases in motivated behavior searching for and 'wanting' the missing reward. Alternatively, the increases in [DA] could signal an aversive motivational component of the reward omission (Amsel, 1958, 1962; Papini and Dudley, 1997; Faure et al., 2008; Richard and Berridge, 2011).

Sub-second DA transmission measured with FSCV

Phasic changes in DA release have been proposed to signal reward prediction errors. Specifically, when an outcome is better than anticipated, a subset of putative midbrain DA neurons reveal a phasic burst of firing; when an outcome is worse than expected, these neurons reveal a brief pause in tonic firing (Schultz et al., 1997; Roesch et al., 2007). Based upon these electrophysiology studies, phasic increases in DA release signaling positive prediction errors and phasic decreases in [DA] signaling negative prediction errors were hypothesized to occur in terminal regions (Schultz, 2002). While much research has focused on such phasic increases in DA release, fewer studies have

examined the effects of brief decreases in DA release, although decreases may be as important, behaviorally, as increases (Goto et al., 2007). Since electrophysiology cannot measure the amount of neurotransmitter release, rapid changes in [DA] needed to be measured at terminal regions in order to understand the resulting changes in neurotransmitter release from previously reported changes in the firing rate of DA neurons. Specifically, during the reward omission manipulation in the utilized behavioral paradigm, a brief decrease in [DA] was expected to occur; to test if this actually happens in the NAc, we utilized FSCV.

Consistent with the reward prediction hypothesis, a decrease in [DA] was detected during the omission of an expected reward during the first block of reward omission (Fig. 4.3Bi.). Once subjects learned about the extinguished lever and reduced responding on this lever (Fig. 4.3A), a significant decrease was no longer detected (block 2; Fig. 4.3Bii.). By the last block of the test session, subjects rarely pressed on the non-rewarded lever. On the few trials in which they did press the non-rewarded lever, a small decrease in [DA] was detected. These few trials could be due to mistakenly pressing the non-rewarded lever; alternatively, subjects could be testing if that lever was still non-rewarded or if the contingencies changed again. Additionally, as subjects gained experience with the non-rewarded lever, cue-evoked DA to the cue light signaling non-rewarded trials decreased, relative to cue-evoked DA for the rewarded option (Fig. 4.4).

Alternative interpretations and challenges to the reward prediction theory of DA neurons

While the theory that midbrain DA neurons signal reward prediction errors has gained strong momentum in the fields of behavioral and computational neuroscience, there is evidence that DA transmission may mediate the attribution of incentive salience

to cues rather than learning the predictive outcomes that the cues signal. Kent Berridge and Terry Robinson have pioneered the theory that DA signals the incentive salience, rather than the learned value or hedonic properties, of stimuli (Robinson and Berridge, 1993; Berridge and Robinson, 1998; Berridge et al., 2009). Indeed, there is evidence that DA is not critical for all forms of learning (Robinson et al., 2005); instead, DA is necessary for endowing reward predictive cues with incentive value and for exhibiting behaviors indicative of cues being incentive stimuli (Flagel et al., 2011; Saunders and Robinson, 2012).

Additionally, a study utilizing FSCV in subjects who differed in their propensity to attribute incentive salience to cues found different DA transmission dynamics in the NAc of these two groups (Flagel et al., 2011). Specifically, in subjects that tend to attribute incentive value to reward predictive cues (termed "sign trackers"), cue-evoked DA release to the conditioned stimulus developed over the course of sessions and became greater than DA release to the unconditioned stimulus. However, these changes did not emerge in rats for which reward predictive cues do not function as incentive stimuli ("goal trackers"). Importantly, it has been well documented that both sign trackers and goal trackers learn the associative relationship between the cue and reward (Yager and Robinson, 2013; Robinson et al., 2014), so difference in DA transmission between the two groups cannot be attributed to a deficit in learning. The DA transmission dynamics in the NAc core are consistent with the reward prediction error theory in sign trackers but not goal trackers. Because both groups learned the predictive nature of the cues and exhibited conditioned responses, if DA transmission in the NAc signals learning, then DA transmission should have been equivalent in both groups. Since DA release differed

between these two subpopulations of rats, this study suggests that DA transmission in the NAc core facilitates the attribution of incentive salience to reward predictive cues and not mere associative learning.

In light of this study, the changes in DA transmission detected in our experiments could represent changes in the incentive value of reward predictive cues. Robust cue-evoked DA was observed in the NAc of our subjects (Fig. 4.2, 4.4), similar to the sign trackers in the aforementioned study (Flagel et al., 2011). The observed decreases in [DA] during the omission of an expected reward could represent inconsistencies between the actual outcome and the expected outcome based upon the incentive stimulus. Additionally, the reduced cue-evoked DA to the non-rewarded cue could signal the loss of incentive value of the non-reinforced cue. Together, our results are consistent with the interpretation that DA is more of a motivational neurotransmitter that attributes incentive value to cues rather than a learning neurotransmitter.

A noteworthy point of divergence between our FSCV data and the classical electrophysiology studies is the time course in which the pause in neuronal firing and the decrease in [DA], respectively, occur. The pauses in firing detected by electrophysiology of midbrain DA neurons in primates have been seen to be temporally locked to the exact moment that the reward was omitted, often within ~100 ms (Schultz et al., 1997). The decrease in [DA] in the NAc that we detect with FSCV reveals a slower, longer-lasting change (Fig. 4.3B), which is consistent with another very recent FSCV study (Hart et al., 2014).

There are a few possible explanations to reconcile this difference in results between FSCV and electrophysiology studies. First, since the utilized behavioral paradigm normally delivers the reward 2 sec after the operant response is completed and the levers recess, the time-locked nature of the decrease in [DA] would be dependent on the rats' abilities to accurately and consistently time exactly 2 sec with millisecond precision. On trials in which rats responded on the non-rewarded lever, no cues indicated precisely 2 sec after the completion of the operant response, other than the passage of time. Therefore, if their timing of 2 sec was even off by a few hundred milliseconds, the decrease in [DA] could occur at a different time. Indeed, based upon behavioral observations of subjects completing this task, we know that they generally approach the food port immediately after the lever recess, not waiting 2 sec until the reward is delivered. Second, a closer examination of the raster plots depicting the pauses in neuronal firing during the omission of an expected reward reveals that the decrease in firing does not always occur during the exact same time point during each trial; indeed, consistent with our FSCV results, the pause in neuronal firing can vary and last up to 0.5-0.75 seconds (Schultz, 2002; Roesch et al., 2007). Third, electrophysiology experiments recording from midbrain DA neurons cannot capture terminal modulation of DA release from other neurons. Therefore, it is possible that inputs from other areas are modulating DA release irrespective of the firing rate of midbrain DA neurons. Additionally, differing densities of the DA transporter can cause differences in extra-cellular [DA] (Wickens et al., 2007; Humphries and Prescott, 2010). Fourth, we now know that the classical studies revealing reward prediction errors through the firing patterns of DA neurons were only recording from a subset of DA neurons, and the midbrain population

of DA neurons is much more heterogeneous than was thought 10 years ago (Ungless et al., 2004; Margolis et al., 2006a; Lammel et al., 2008; Brischoux et al., 2009; Ungless and Grace, 2012). For example, an electrophysiology study examining reward prediction errors in the DA neurons of rats recorded from 258 neurons in the VTA and substantia nigra but only 36 of these neurons were classified as dopaminergic based upon traditional electrophysiological criteria of DA neurons (Roesch et al., 2007), although anatomical studies reveal that approximately 55% of VTA neurons and at least 88% of substantia nigra neurons are dopaminergic (Margolis et al., 2006a). Therefore, it is possible that some of the firing patterns affecting DA release in terminal regions have not been fully captured by electrophysiology.

Similarly, it has been suggested that the temporal precision of the phasic changes in DA neuronal firing occur too rapidly to actually signal reward prediction errors (Redgrave and Gurney, 2006; Redgrave et al., 2008). The phasic changes in putative midbrain DA neuronal firing can occur very rapidly, within ~100 ms (Schultz, 2002). For these changes to be signaling violations of expectation, the stimulus or outcome must be identified by the brain in order to compute whether a violation of expectation has occurred. Peter Redgrave has drawn attention to the fact that the identity of visual stimuli (which are primarily used in electrophysiology and reward prediction error experiments) may not be well discriminated by the brain until ~80-160 ms after stimuli onset (Schall, 2003; Redgrave et al., 2008). However, a direct projection from the superior colliculus to the VTA and substantia nigra pars compacta has been discovered, and this projection is functionally important for the ability of visual stimuli to evoked phasic increases in

putative DA neuronal firing (Comoli et al., 2003; Dommett et al., 2005; Overton et al., 2014).

Based upon these findings, it has been proposed that DA neurons may reinforce the discovery of unpredicted sensory events rather than signaling specific prediction errors (Redgrave et al., 2008). And, indeed, there is evidence that some DA neurons fire to salient sensory events, regardless of their valence or predictive value (Matsumoto and Hikosaka, 2009; Bromberg-Martin et al., 2010b). However, this does not necessarily preclude the possibility that some DA neurons may signal reward prediction errors based upon the variations in the timing of the responses explained in the previous paragraph. Indeed, recent experiments demonstrate that a subpopulation of DA neurons (primarily in the VTA and ventromedial substantia nigra pars compacta) appear to signal reward prediction errors while more dorsolateral DA neurons (primarily in the substantia nigra pars compacta) respond to salient stimuli necessary for completing a working memory task (Matsumoto and Takada, 2013).

Another growing problem for the reward prediction error hypothesis of DA neurons is that there is accumulating evidence that many DA neurons do not behave according to the classical reward prediction error theory proposed by Schultz. Indeed, some DA neurons increase firing to aversive events (Brischoux et al., 2009; Matsumoto and Hikosaka, 2009; Lammel et al., 2011) and, as stated above, some DA neurons fire to novel stimuli, regardless of the valence that these stimuli signal (Horvitz, 2000; Matsumoto and Hikosaka, 2009; Overton et al., 2014).

Converging evidence suggests that there are different subpopulations of DA neurons that function differently, signaling different things. Indeed, there is growing consensus that midbrain DA neurons are not homogenous, as was once thought (Lammel et al., 2008; Lammel et al., 2011). Therefore, Okihide Hikosaka and colleagues have suggested that there are at least three types of midbrain DA neurons: DA neurons that signal motivational salience, neurons that signal motivational value, and some that provide alerting signals (Bromberg-Martin et al., 2010b). Furthermore, there is evidence that the same neurons can mediate more than one of the aforementioned roles, such as signaling valence and salience (Bromberg-Martin et al., 2010a).

Given the diverse projections (Ikemoto, 2007; Lammel et al., 2011) and the varying electrophysiological characteristics of DA midbrain DA neurons (Margolis et al., 2006a; Lammel et al., 2008), it is plausible that DA neuronal responses to environmental events can be very diverse. Additionally, environmental factors (such as stress) can affect the proportions of the NAc in which glutamate transmission elicits appetitive versus aversive behaviors (Reynolds and Berridge, 2008), and these effects are dependent on DA transmission (Faure et al., 2008; Richard and Berridge, 2011).

Much debate has centered on the role of DA transmission in motivated behavior (Berridge and Robinson, 1998; Wise, 2004; Salamone and Correa, 2012). However, when taken together, the converging results from the numerous experiments trying to elucidate the role of DA suggest that DA does not do any one particular thing. Rather, DA neurons receiving various inputs and projecting to varying forebrain regions likely convey different signals depending upon the environment and motivational state of the animal.

Behavioral preferences prompted by the omission of an expected reward are mediated by D2-like receptors in the NAc

Electrophysiology studies of classical midbrain DA neurons have revealed that this subpopulation of neurons exhibit a phasic reduction in firing rate when an expected reward is omitted, and in chapter 4, we demonstrated that there is a decrease in [DA] during this event. Modeling data has shown that phasic decreases in [DA] reduce D1 and D2 receptor occupancy to 0% (Dreyer et al., 2010)[2]. However, because of the differing affinities of D1 and D2 receptors to DA (Richfield et al., 1989; Dreyer et al., 2010), decreases in [DA] have been hypothesized to preferentially affect D2 receptors (Frank et al., 2004; Frank, 2005; Bromberg-Martin et al., 2010b; Dreyer et al., 2010). Specifically, baseline occupancy of D1 receptors during regular tonic activity is estimated to be ~ 3.5%, whereas D2 occupancy is ~75%; therefore, a reduction in [DA] due to pauses in dopaminergic neuronal firing (as would happen during reward omission) would have greater effects at D2 receptors (Dreyer et al., 2010).

If reductions in [DA] to the omission of an expected reward are critical for altering behavior in response to the extinguished lever, and if these reductions in [DA] are signaled through reductions in binding at post-synaptic D2 receptors, then preventing the decrease in occupancy of DA at D2 receptors through administration of a D2-like agonist would be expected to attenuate the effects of reward omission. We tested this hypothesis by site-specifically administering D1-like and D2-like DA receptor agonists and antagonists prior to subjects experiencing reward omission. We found that only the

[2] While the modeling data was primarily focused on DA receptors in the dorsal striatum, D1 and D2 receptors in the NAc core have similar affinities to DA as they do in the dorsal striatum (Richfield et al., 1989).

D2-like agonist dose dependently attenuated the behavioral preference for the rewarded option (Fig. 5.2). Furthermore, any general locomotor effects of the drug are not responsible for the lack of behavioral preference for the rewarded option (Fig. 5.5). These results support the hypothesis that phasic reductions in [DA] acting via D2-like receptors in the NAc are important for altering behavioral strategy following the omission of an expected reward.

Additionally, these data support the theory that striatal neurons containing D2 receptors, which are part of the indirect pathway, suppress actions and modify future behavior away from aversive outcomes (Bromberg-Martin et al., 2010b; Hikida et al., 2010; Kravitz and Kreitzer, 2012; Kravitz et al., 2012). The majority of neurons in the striatum (~95-97%) are medium spiny neurons (Kemp and Powell, 1971; Tepper and Bolam, 2004; Matamales et al., 2009), and these medium spiny neurons in the striatum primarily express either D1 or D2 receptors, although co-localization of both these receptor types is found on some neurons (Matamales et al., 2009). D1-expressing neurons are generally part of the direct pathway, whereas D2-expressing neurons are predominantly part of the indirect pathway (Humphries and Prescott, 2010; Gerfen and Surmeier, 2011). In the dorsal striatum, the direct pathway projects "directly" to the internal globus pallidus and substantia nigra pars reticulada, whereas the indirect pathway detours through the external globus pallidus and subthalamic nucleus before arriving in the substantia nigra (Berke and Hyman, 2000; Kravitz and Kreitzer, 2012).

While much research has focused on the direct and indirect pathways of the dorsal striatum due to its roles in movement and Parkinson's disease, less has focused on the ventral striatum (Humphries and Prescott, 2010). The core subregion of the NAc is

anatomically similar to the dorsal striatum with projections to the dorsal portion of the substantia nigra pars reticulate (Deniau et al., 1994); however, a number of neurons in the NAc core project to the ventral pallidum (Maurice et al., 1997). Indeed, many D1-expressing neurons in the NAc core project directly to midbrain DA neurons (Deniau et al., 1994; Lu et al., 1997), while a proportion of D1-expressing neurons project to the ventral pallidum (Robertson and Jian, 1995; Zhou et al., 2003). D2-expressing neurons in the NAc core project to the ventral pallidum (Lu et al., 1997). Additionally, since DA transmission dynamics in the NAc core are similar to in the dorsal striatum (Brown et al., 2011), the indirect pathway from the NAc core may serve a similar role in directing motivated behavior away from a negative outcome, as the indirect pathway exiting the dorsal striatum does (Frank, 2005; Hikida et al., 2010; Kravitz and Kreitzer, 2012; Kravitz et al., 2012; Freeze et al., 2013). Supporting this idea, other recent studies have shown that decreased phasic DA transmission in the NAc core is correlated with aversive cues and freezing behavior (Badrinarayan et al., 2012; Oleson et al., 2012).

Reconciling simultaneous increase and decreases in [DA]

In chapter 3 we revealed that when an expected reward is omitted, a tonic increase in [DA] occurs in the NAc measured by microdialysis, while in chapter 4, we emphasize the importance of phasic decreases in DA transmission during reward omission detected with FSCV. Reporting increases and decreases in DA transmission during the same behavioral manipulation may initially appear contradictory, but this need not be the case. Indeed, DA transmission is complex acting on multiple timescales (Grace et al., 2007; Schultz, 2007). In Fig. 6.1, I propose a model synthesizing the FSCV and microdialysis

data in which increases and decreases in [DA] are both occurring and behaviorally relevant for distinct reasons.

Overall, there is an increase in tonic levels of [DA] in the NAc; however, embedded within this change in basal levels are phasic increases and decreases in DA release (Fig. 6.1A). The longer lasting changes in tonic levels of DA may signal important changes in the environmental state or motivational state of the animal. For example, when [DA] increases during reward omission, this tonic increase in [DA] could motivate and invigorate the animal to search for ways to obtain food. This is supported by the fact that these levels of DA correlate to motivational vigor and exploration (Figs. 3.2-3.3). In contrast, phasic, sub-second changes in [DA] often appear to signal temporally precise, behaviorally relevant events as opposed to longer lasting states. For example, phasic changes in [DA] occur to cues (cue-evoked DA) and adjust to the changing value of those cues (Figs. 4.4-4.5). Additionally, small phasic increases and decreases in [DA] occur to reward delivery or reward omission, respectively (Fig. 4.3).

Phasic increases and decreases in DA likely are signaled through different mechanisms and DA receptors on striatal medium spiny neurons. Much of the differences in signaling have to do with the different affinities that D1 and D2 receptors have for DA (Richfield et al., 1989). Models of D1 and D2 receptor occupancy in relation to tonic and phasic DA release events have been generated based upon experimental data (Dreyer et al., 2010). As stated earlier, low-affinity D1 receptors only are ~3.5% occupied, whereas high-affinity D2 receptors are ~75% occupied, during baseline, tonic levels of DA in the striatum (Dreyer et al., 2010). Phasic bursts of DA release (which can exceed 1 μM) briefly saturate the low-affinity D1 receptors (for a few

milliseconds) and then only continue to stimulate D1 receptors closer to the DA release sites (for ~60 milliseconds). Additionally, phasic bursts increase binding of DA at D2 receptors (to >95% occupancy), but since the binding of these receptors is already high (at least 75% occupancy), the relative change in binding at D2 receptors in not nearly as drastic as the changes occurring at D1 receptors. Therefore, phasic increases in DA release are thought to be signaled through increased stimulation of D1 receptors (Fig. 6.1B).

In contrast, brief decreases in [DA] (resulting from pauses in DA release) can reduce DA binding to ~0% at D1 and D2 receptors. Although there is a reduction in binding at both D1 and D2 receptors, since DA occupancy at D1 receptors is already quite low (~3.5%), the reduction would have more profound effects at D2 receptors; therefore, decreases in [DA] are hypothesized to be signaled through decreased occupancy of DA at D2 receptors (Fig. 6.1C). This theory is supported by the fact that phasic decreases in [DA] are detected during reward omission (Fig. 4.3) and holding D2 receptor occupancy high, via an agonist, prevents a reduction in responding on the non-rewarded lever (Fig. 5.2). Additionally, reducing [DA] by DA depletion (from reserpine) decreases spine density, dendritic length, and dendritic branching in striatopallidal neurons (predominantly D2-expressing neurons) but not striatonigral neurons (predominantly D1-expressing neurons), supporting the idea that reductions in DA are signaled preferentially through D2 receptors (Day et al., 2006; Surmeier et al., 2007).

So far we have shown that decreases in DA are likely signaled through a reduction in occupancy of DA at D2 receptors. Based upon the transmission dynamics of DA and the affinity of D2 receptors for DA, decreases in [DA] are likely detected by the

high affinity D2 receptors during normal, tonic levels of extracellular DA (Dreyer et al., 2010). However, in chapter 3 we saw that tonic levels of DA increase to about 200% of baseline levels (Fig. 3.2A). I hypothesize that given what we know about DA transmission and D2 receptors, having higher levels of tonic [DA] in the NAc may actually result in more effective signaling of the phasic decreases in DA release. Extracellular levels of [DA] in ventral striatum are generally low, and these low levels can be advantageous for contrasting certain tonic and phasic signals (Zhang et al., 2009). Indeed, a stimulus train mimicking phasic firing patterns causes greater DA release in the NAc core and shell, where tonic levels of [DA] are lower, compared to the dorsolateral striatum, where tonic levels of [DA] are higher (Keck et al., 2002; Fadda et al., 2005; Zhang et al., 2009). In other words, the contrast between low basal levels and high phasic release events can be beneficial in signaling behaviorally relevant events through phasic increases in DA release. However, if decreases in [DA] are signaled through reductions in activity at D2 receptors (Bromberg-Martin et al., 2010b; Porter-Stransky et al., 2013), then the effects of the reduction in DA at D2 receptors would be most pronounced when basal levels of DA are high.

Behaviorally, given the different hypothesized roles of tonic and phasic DA transmission in the NAc, higher tonic levels of DA may be particularly adaptive when phasic decreases related to negative events occur. As stated above, when environmental contingencies change, it can be adaptive to have higher levels of tonic [DA] to invigorate motivated behavior and bias behavior towards exploring the new environment and employing new behavioral strategies (Beeler et al., 2010; Beeler et al., 2012; Humphries et al., 2012; Baudonnat et al., 2013; Beeler et al., 2014). When testing new behavioral

responses and strategies in this invigorated state, undesirable outcomes may be signaled more clearly through decreases in DA transmission at D2 receptors when tonic levels of [DA] are high.

Furthermore, decreased binding of DA at D2 receptors, as would happen during the omission of an expected reward, has been shown to facilitate prefrontal cortical inputs into the nucleus accumbens (Goto and Grace, 2005). Additionally, stimulation of prefrontal cortical neurons can attenuate glutamatergic inputs from the hippocampus and thalamus into the ventral striatum, thus biasing the ventral striatum toward incoming information from the prefrontal cortex (Calhoon and O'Donnell, 2013), and the prefrontal cortex has been shown to be important for mediating certain forms of behavioral flexibility (Ragozzino et al., 2003; Ragozzino, 2007; Floresco, 2013). Together, reductions in D2 receptor activation could cause synaptic plasticity and thereby function as a neural mechanism for altering response strategy and mediating behavioral flexibility (Meck and Benson, 2002; Goto and Grace, 2005).

An important caveat not yet addressed is that while D2 receptors generally have a higher affinity for DA than D1 receptors, D2 receptors can switch between a high ($D2^{High}$) and low ($D2^{Low}$) affinity state (Richfield et al., 1989; Seeman, 2011; van Wieringen et al., 2013). When in the $D2^{High}$ state, DA would more readily bind to the D2 receptors than when the receptors are in the $D2^{Low}$ state; thus, the $D2^{High}$ state is thought to be the functional state of the receptor (George et al., 1985). Therefore, the estimations provided in the above model (Dreyer et al., 2010) could be affected by changes in the affinity states of D2 receptors. While D2 receptors have previously been shown to primarily be in the $D2^{High}$ state (Richfield et al., 1989), other studies have more recently shown the

percentage of $D2^{High}$ receptors to be ~20% (Seeman et al., 2005; Seeman, 2009). Additionally, pharmacological manipulations can substantially alter the proportion of $D2^{Low}$ versus $D2^{High}$, suggesting that a number of D2 receptors are in the $D2^{Low}$ state. For example, cocaine self-administration can cause a significant increase (~150%) in the proportion of $D2^{High}$ state receptors without altering the overall number of D2 receptors (Briand et al., 2008). However, it is important to remember that most studies examining the affinity states of D2 receptors have been done in homogenized brain tissues; therefore, the proportion of $D2^{High}$ versus $D2^{Low}$ receptors *in vivo* during specific motivated behaviors has yet to be elucidated (Seeman, 2011).

Of particular relevance to the current experiments, individual variation has been observed in the exploratory behavior of rats. Interestingly, rats that are classified as high explores have higher extracellular levels of [DA] in the dorsal striatum (measured by microdialysis) and a higher proportion of $D2^{High}$ state receptors than low explorer rats (Alttoa et al., 2009). Additionally, these high explorer rats also show greater increases in striatal [DA] to amphetamine than low explorer rats (Alttoa et al., 2009), demonstrating that while they may have higher baseline levels of striatal [DA], substantial increases in [DA] are still possible in these rats. While much is still unknown about changes in affinity states of D2 receptors in relation to behavioral events, and thus far the data are only correlational, these experiments raise the intriguing possibility that the increased exploration observed during reward omission, along with the corresponding increases in [DA] in the NAc, may relate to a shift in D2 receptors to the $D2^{High}$ state. Having a greater percentage of D2 receptors in the $D2^{High}$ state could potentially be a neural mechanism for signaling increased exploratory behaviors. Furthermore, the phasic

decreases in [DA] that would occur to non-reinforced actions could cause more substantial reductions in D2 receptor binding when a greater number of receptors are in this $D2^{High}$ state. Currently, these ideas are still speculative, so future studies are needed to test these hypotheses.

Conclusion

In this dissertation, I have revealed multiple, distinct DA transmission patterns in the NAc core during the omission of an expected reward in an operant foraging task. Specifically, tonic levels of [DA] measured by microdialysis increase to ~200% of baseline levels and correspond to increases in motivational vigor and exploration of alternative strategies. Concurrently, sub-second, phasic changes in DA release are occurring to behaviorally relevant events. Specifically, when an expected reward is not received, a phasic decrease in [DA] occurs, which could be facilitating learning about the non-rewarded option or reducing the incentive value of that option. As subjects learn about the non-rewarded lever, cue-evoked DA to the non-rewarded option decreases, and, behaviorally, subjects develop a strong preference for the rewarded option and avoid the non-rewarded option. Finally, this alteration in behavioral preference is mediated by a reduction in occupancy of D2 receptors in the NAc core. Together, this series of experiments provides insight into the complex transmission dynamics of DA in the NAc in relation to flexible behavior during reductions in reward availability.

Future Directions

While techniques like microdialysis and FSCV provide insight into the neurotransmission dynamics of DA, they do not reveal causal roles of neurotransmission on behavior. The ability of a D2 agonist in the NAc to cause subjects to keep responding

on the non-rewarded lever suggests that decreases in [DA] are necessary for altering behavioral strategy; however, the D2 agonist is presumably acting continually on receptors throughout the session. To determine if the temporally precise phasic decrease during the omission of the expected reward is the key to modifying behavioral response, a technique that can alter DA transmission with sub-second precision is necessary.

Optogenetics is a technique utilizing light-sensitive proteins, often in combination with transgenic mice or rats, that can facilitate the manipulation of specific neurotransmitter systems with millisecond precision (Deisseroth, 2011; Fenno et al., 2011; Witten et al., 2011; Johansen et al., 2012). A recent study suggests that preventing phasic decreases in [DA] during Pavlovian extinction attenuates learning (Steinberg et al., 2013). To test whether the phasic decrease in [DA] that we observed with FSCV is necessary for altering behavioral strategy away from the non-rewarded option, in a follow up study, we are delivering a brief optical stimulation to dopaminergic neurons in the VTA to "fill in the dip" of [DA] (Frank, 2005) after subjects respond on the non-rewarded levers, causing [DA] on the non-rewarded trials to be similar to that on the rewarded trials (Fig. 4.3B). While the experiment is not yet complete, we have pilot data supporting this hypothesis. After subjects complete the operant response on the non-rewarded lever, blue light is delivered to the VTA. Optical stimulation of the VTA of TH-Cre+ rats expressing channelrhodopsin (Fig. 6.2A) can attenuate the behavioral preference for the rewarded option (Fig. 6.2B). And, high levels of optic stimulation actually appear to create a preference for the non-rewarded lever (Fig.6.2B). While these data are preliminary, they suggest that the phasic decrease in [DA] is functionally necessary to alter reward seeking behavior away from the non-rewarded option.

Similarly, optogenetics may be useful to elucidate if the tonic increases in [DA] observed using microdialysis during reward omission mediate changes in response vigor and exploration. First, we would need to determine the appropriate parameters of optic stimulation that cause the same magnitude of an increase in [DA] measured using microdialysis. Then, tonic levels of [DA] could be altered optogenetically and changes in behavioral response vigor and exploration recorded.

Figures

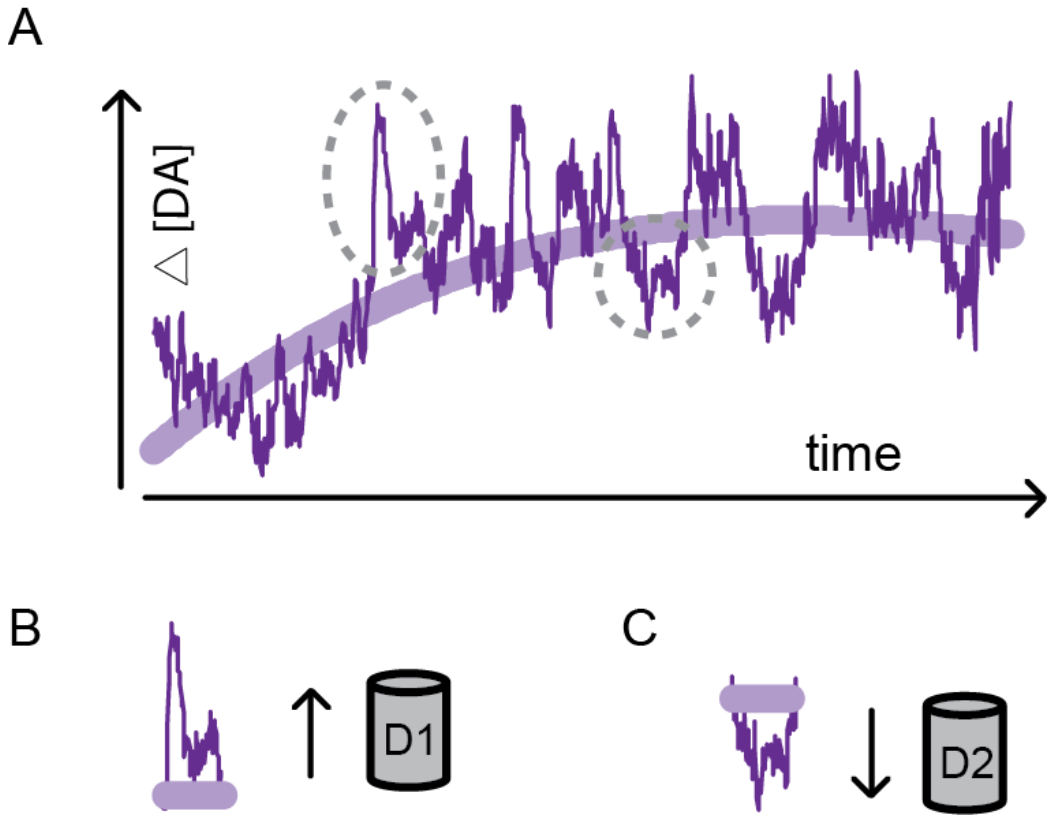

Figure 6.1. Model of DA transmission in the NAc core during unexpected reward omission. **A)** Distinct phasic (purple) and tonic (violet) components of DA transmission can occur simultaneously. Examples of phasic increases and decreases (circled in gray, dotted lines), which often occur in relation to behaviorally relevant events, can ride on top of slower changing basal levels of extracellular [DA]. **B-C)** Increases and decreases in [DA] likely signaled via different DA receptors. Specifically, phasic increases in DA can activate low affinity D1 receptors (B), whereas decreases in [DA] are likely transmitted through a reduction in DA binding at high affinity D2 receptors.

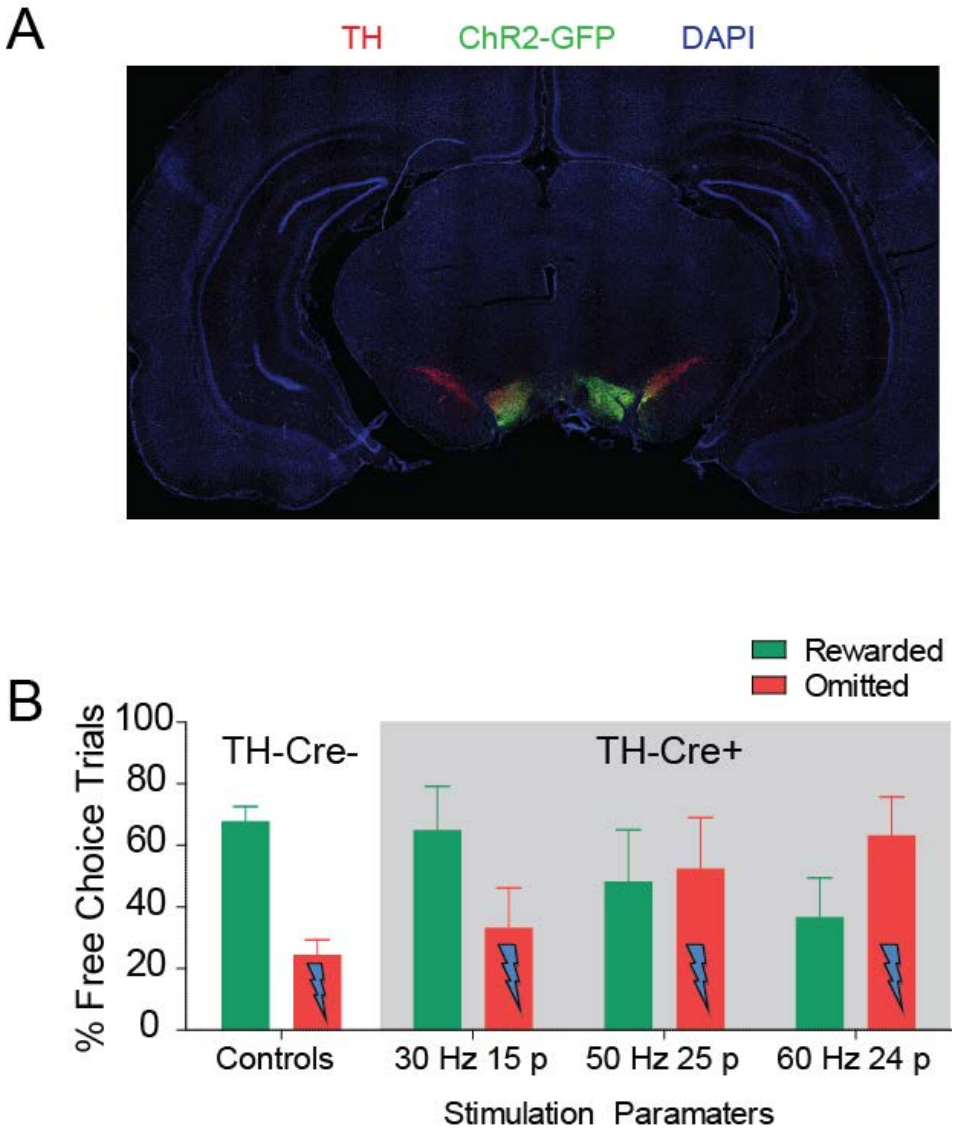

Figure 6.2. Optogenetic activation of DA neurons during reward omission attenuates the behavioral preference for the rewarded option. **A)** A representative image taken from a TH-Cre+ rat injected with a Cre-dependent channelrhodopsin-containing viral vector. **B)** Behavioral responding on free choice trials. While low levels of stimulation in TH-Cre-+ rats do not impair the behavioral preference for the rewarded option, medium levels attenuate the behavioral preference. Additionally, high levels of stimulation actually cause greater responding on the non-rewarded lever. These effects are not seen in wild-type littermates (TH-Cre-). *N* ~ 3-4 per group.

REFERENCES

Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006) Prediction error as a linear function of reward probability is coded in human nucleus accumbens. Neuroimage 31:790-795.

Adelman HM, Maatsch JL (1956) Learning and extinction based upon frustration, food reward, and exploratory tendency. Journal of Experimental Psychology 52:311.

Agnati LF, Guidolin D, Guescini M, Genedani S, Fuxe K (2010) Understanding wiring and volume transmission. Brain Research Reviews 64:137-159.

Aharonovich E, Hasin DS, Brooks AC, Liu X, Bisaga A, Nunes EV (2006) Cognitive deficits predict low treatment retention in cocaine dependent patients. Drug and Alcohol Dependence 81:313-322.

Ahn S, Phillips AG (2007) Dopamine efflux in the nucleus accumbens during within-session extinction, outcome-dependent, and habit-based instrumental responding for food reward. Psychopharmacology 191:641-651.

Alttoa A, Seeman P, Koiv K, Eller M, Harro J (2009) Rats with persistently high exploratory activity have both higher extracellular dopamine levels and higher proportion of D 2High receptors in the striatum. Synapse 63:443-446.

Amsel A (1958) The role of frustrative nonreward in noncontinuous reward situations. Psychological Bulletin 55:102.

Amsel A (1962) Frustrative nonreward in partial reinforcement and discrimination learning: Some recent history and a theoretical extension. Psychological Review; Psychological Review 69:306.

Annett LE, McGregor A, Robbins TW (1989) The effects of ibotenic acid lesions of the nucleus accumbens on spatial learning and extinction in the rat. Behav Brain Res 31:231-242.

Anstrom KK, Miczek KA, Budygin EA (2009) Increased phasic dopamine signaling in the mesolimbic pathway during social defeat in rats. Neuroscience 161:3-12.

Aragona BJ (2011) The regional specificity of rapid actions of cocaine. Nature Reviews Neuroscience 12:700-700.

Aragona BJ, Wang Z (2009) Dopamine regulation of social choice in a monogamous rodent species. Frontiers in behavioral neuroscience 3.

Aragona BJ, Cleaveland NA, Stuber GD, Day JJ, Carelli RM, Wightman RM (2008) Preferential enhancement of dopamine transmission within the nucleus accumbens shell by cocaine is attributable to a direct increase in phasic dopamine release events. J Neurosci 28:8821-8831.

Aragona BJ, Day JJ, Roitman MF, Cleaveland NA, Wightman RM, Carelli RM (2009) Regional specificity in the real-time development of phasic dopamine transmission patterns during acquisition of a cue-cocaine association in rats. Eur J Neurosci 30:1889-1899.

Aragona BJ, Liu Y, Yu YJ, Curtis JT, Detwiler JM, Insel TR, Wang Z (2006) Nucleus accumbens dopamine differentially mediates the formation and maintenance of monogamous pair bonds. Nat Neurosci 9:133-139.

Arbuthnott GW, Wickens J (2007) Space, time and dopamine. Trends Neurosci 30:62-69.

Badiani A, Morano MI, Akil H, Robinson TE (1995) Circulating adrenal hormones are not necessary for the development of sensitization to the psychomotor activating effects of amphetamine. Brain research 673:13-24.

Badrinarayan A, Wescott SA, Vander Weele CM, Saunders BT, Couturier BE, Maren S, Aragona BJ (2012) Aversive Stimuli Differentially Modulate Real-Time Dopamine Transmission Dynamics within the Nucleus Accumbens Core and Shell. The Journal of neuroscience 32:15779-15790.

Baker TW, Weisman RG, Beninger RJ (2012) Reinforcer devaluation by extinction depends on the food restriction protocol. Behavioural processes.

Baldo BA, Sadeghian K, Basso AM, Kelley AE (2002) Effects of selective dopamine D1 or D2 receptor blockade within nucleus accumbens subregions on ingestive behavior and associated motor activity. Behavioural brain research 137:165-177.

Baudonnat M, Huber A, David V, Walton ME (2013) Heads for learning, tails for memory: reward, reinforcement and a role of dopamine in determining behavioral relevance across multiple timescales. Frontiers in Neuroscience 7.

Becker JB (2009) Sexual differentiation of motivation: a novel mechanism? Horm Behav 55:646-654.

Becker JB, Taylor JR (2008) Sex differences in motivation. In: Sex differences in the brain: from genes to behavior (Becker JB, ed), pp 177-199: Oxford University Press.

Becker JB, Arnold AP, Berkley KJ, Blaustein JD, Eckel LA, Hampson E, Herman JP, Marts S, Sadee W, Steiner M, Taylor J, Young E (2005) Strategies and methods for research on sex differences in brain and behavior. Endocrinology 146:1650-1673.

Beeler JA (2012) Thorndike's Law 2.0: Dopamine and the Regulation of Thrift. Frontiers in Neuroscience 6.

Beeler JA, Frazier CR, Zhuang X (2012) Putting desire on a budget: dopamine and energy expenditure, reconciling reward and resources. Frontiers in integrative neuroscience 6.

Beeler JA, Daw ND, Frazier CR, Zhuang X (2010) Tonic dopamine modulates exploitation of reward learning. Frontiers in behavioral neuroscience 4:170.

Beeler JA, Cools R, Luciana M, Ostlund SB, Petzinger G (2014) A kinder, gentler dopamine… highlighting dopamine's role in behavioral flexibility. Frontiers in Neuroscience 8.

Beery AK, Zucker I (2011) Sex bias in neuroscience and biomedical research. Neurosci Biobehav Rev 35:565-572.

Benedix J (1993) Area-restricted search by the plains pocket gopher (Geomys bursarius) in tallgrass prairie habitat. Behavioral Ecology 4:318-324.

Berke JD, Hyman SE (2000) Addiction, dopamine, and the molecular mechanisms of memory. Neuron 25:515-532.

Berridge KC (2012) From prediction error to incentive salience: mesolimbic computation of reward motivation. European Journal of Neuroscience 35:1124-1143.

Berridge KC, Robinson TE (1998) What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Research Reviews 28:309-369.
Berridge KC, Robinson TE, Aldridge JW (2009) Dissecting components of reward: 'liking', 'wanting', and learning. Curr Opin Pharmacol 9:65-73.
Bonci A, Malenka RC (1999) Properties and plasticity of excitatory synapses on dopaminergic and GABAergic cells in the ventral tegmental area. The Journal of neuroscience 19:3723-3730.
Borgkvist A, Mrejeru A, Sulzer D (2011) Multiple Personalities in the Ventral Tegmental Area. Neuron 70:803-805.
Boschen SL, Wietzikoski EC, Winn P, Cunha CD (2011) The role of nucleus accumbens and dorsolateral striatal D2 receptors in active avoidance conditioning. Neurobiology of learning and memory 96:254-262.
Boye SM, Grant RJ, Clarke P (2001) Disruption of dopaminergic neurotransmission in nucleus accumbens core inhibits the locomotor stimulant effects of nicotine and D-amphetamine in rats. Neuropharmacology 40:792-805.
Briand LA, Flagel SB, Seeman P, Robinson TE (2008) Cocaine self-administration produces a persistent increase in dopamine D2 High receptors. Eur Neuropsychopharmacol 18:551-556.
Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009) Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. Proc Natl Acad Sci U S A 106:4894-4899.
Britt JP, McGehee DS (2008) Presynaptic opioid and nicotinic receptor modulation of dopamine overflow in the nucleus accumbens. The Journal of neuroscience 28:1672.
Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010a) Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. Neuron 67:144-155.
Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010b) Dopamine in motivational control: rewarding, aversive, and alerting. Neuron 68:815-834.
Brown HD, McCutcheon JE, Cone JJ, Ragozzino ME, Roitman MF (2011) Primary food reward and reward-predictive stimuli evoke different patterns of phasic dopamine signaling throughout the striatum. Eur J Neurosci.
Brown SD, Peters JC (2004) Hydrogenolysis of [PhBP3]Fe[triple bond]N-p-tolyl: probing the reactivity of an iron imide with H2. J Am Chem Soc 126:4538-4539.
Cachope R, Mateo Y, Mathur BN, Irving J, Wang H-L, Morales M, Lovinger DM, Cheer JF (2012) Selective activation of cholinergic interneurons enhances accumbal phasic dopamine release: setting the tone for reward processing. Cell reports.
Calhoon GG, O'Donnell P (2013) Closing the gate in the limbic striatum: prefrontal suppression of hippocampal and thalamic inputs. Neuron 78:181-190.
Cameron D, Wessendorf M, Williams J (1997) A subset of ventral tegmental area neurons is inhibited by dopamine, 5-hydroxytryptamine and opioids. Neuroscience 77:155-166.
Cannon CM, Palmiter RD (2003) Reward without dopamine. The Journal of neuroscience 23:10827-10831.

Cardinal RN, Pennicott DR, Lakmali C, Robbins TW, Everitt BJ (2001) Impulsive choice induced in rats by lesions of the nucleus accumbens core. Science 292:2499-2501.
Carroll ME, Mach JL, La Nasa RM, Newman JL (2009) Impulsivity as a behavioral measure of withdrawal of orally delivered PCP and nondrug rewards in male and female monkeys. Psychopharmacology 207:85-98.
Choi JE, Vaswani PA, Shadmehr R (2014) Vigor of movements and the cost of time in decision making. The Journal of neuroscience 34:1212-1223.
Clark DA (1980) Age- and Sex-Dependent Foraging Strategies of a Small Mammalian Omnivore. Journal of Animal Ecology 49:549-563.
Clark JJ, Collins AL, Sanford CA, Phillips PE (2013) Dopamine Encoding of Pavlovian Incentive Stimuli Diminishes with Extended Training. The Journal of neuroscience 33:3526-3532.
Cohen JD, McClure SM, Angela JY (2007) Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philosophical Transactions of the Royal Society B: Biological Sciences 362:933-942.
Colivicchi MA, Stefanini C, Freinbichler W, Ballini C, Bianchi L, Tipton KF, Della Corte L (2013) Monitoring Extracellular Amino Acid Neurotransmitters and hROS by In Vivo Microdialysis in Rats: A Practical Approach. In: Microdialysis Techniques in Neuroscience, pp 225-247: Springer.
Colzato LS, Huizinga M, Hommel B (2009) Recreational cocaine polydrug use impairs cognitive flexibility but not working memory. Psychopharmacology 207:225-234.
Comoli E, Coizet V, Boyes J, Bolam JP, Canteras NS, Quirk RH, Overton PG, Redgrave P (2003) A direct projection from superior colliculus to substantia nigra for detecting salient visual events. Nature neuroscience 6:974-980.
Cools R, Altamirano L, D'Esposito M (2006) Reversal learning in Parkinson's disease depends on medication status and outcome valence. Neuropsychologia 44:1663-1673.
Cools R, Barker RA, Sahakian BJ, Robbins TW (2001) Mechanisms of cognitive set flexibility in Parkinson's disease. Brain 124:2503-2512.
Cools R, Lewis SJG, Clark L, Barker RA, Robbins TW (2007) L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. Neuropsychopharmacology 32:180-189.
Corbit LH, Muir JL, Balleine BW (2001) The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. J Neurosci 21:3251-3260.
Cragg SJ (2006) Meaningful silences: how dopamine listens to the ACh pause. Trends in neurosciences 29:125-131.
Crespi LP (1942) Quantitative variation of incentive and performance in the white rat. Am J Psychol 55:467-517.
Crombag H, Mueller H, Browman K, Badiani A, Robinson T (1999) A comparison of two behavioral measures of psychomotor activation following intravenous amphetamine or cocaine: dose-and sensitization-dependent changes. Behavioural Pharmacology 10:205.
Cummings JA, Gowl BA, Westenbroek C, Clinton SM, Akil H, Becker JB, Cummings J, Gowl B, Clinton S (2011) Effects of a selectively bred novelty-seeking phenotype

on the motivation to take cocaine in male and female rats. Biology of sex differences 2:1-10.
Dalla C, Shors TJ (2009) Sex differences in learning processes of classical and operant conditioning. Physiology & Behavior 97:229-238.
Dalley JW, Everitt BJ (2009) Dopamine receptors in the learning, memory and drug reward circuitry. Seminars in Cell & Developmental Biology 20:403-410.
Daly HB (1969a) Learning of a hurdle-jump response to escape cues paired with reduced reward or frustrative nonreward. Journal of Experimental Psychology 79:146.
Daly HB (1969b) Aversive properties of partial and varied reinforcement during runway acquisition. Journal of Experimental Psychology 81:54.
Daly HB (1969c) Is instrumental responding necessary for nonreward following reward to be frustrating? Journal of Experimental Psychology 80:186.
Daly HB (1974) Reinforcing properties of escape from frustration aroused in various learning situations. The psychology of learning and motivation 8:187-231.
Darvas M, Fadok JP, Palmiter RD (2011) Requirement of dopamine signaling in the amygdala and striatum for learning and maintenance of a conditioned avoidance response. Learning & memory 18:136-143.
David HN, Sissaoui K, Abraini JH (2004) Modulation of the locomotor responses induced by D1-like and D2-like dopamine receptor agonists and D-amphetamine by NMDA and non-NMDA glutamate receptor agonists and antagonists in the core of the rat nucleus accumbens. Neuropharmacology 46:179-191.
Davis BA, Clinton SM, Akil H, Becker JB (2008) The effects of novelty-seeking phenotypes and sex differences on acquisition of cocaine self-administration in selectively bred High-Responder and Low-Responder rats. Pharmacology Biochemistry and Behavior 90:331-338.
Daw ND, Touretzky DS (2002) Long-term reward prediction in TD models of the dopamine system. Neural Computation 14:2567-2583.
Daw ND, Doya K (2006) The computational neurobiology of learning and reward. Current opinion in neurobiology 16:199-204.
Daw ND, Kakade S, Dayan P (2002) Opponent interactions between serotonin and dopamine. Neural Networks 15:603-616.
Day JJ, Jones JL, Carelli RM (2011) Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. European Journal of Neuroscience 33:308-321.
Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat Neurosci 10:1020-1028.
Day JJ, Jones JL, Wightman RM, Carelli RM (2010) Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. Biol Psychiatry 68:306-309.
Day M, Wang Z, Ding J, An X, Ingham CA, Shering AF, Wokosin D, Ilijic E, Sun Z, Sampson AR, Mugnaini E, Deutch AY, Sesack SR, Arbuthnott GW, Surmeier DJ (2006) Selective elimination of glutamatergic synapses on striatopallidal neurons in Parkinson disease models. Nature neuroscience 9:251-259.
Dayan P, Berridge KC (2014) Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. Cognitive, Affective, & Behavioral Neuroscience:1-20.

Deisseroth K (2011) Optogenetics. Nature methods 8:26-29.
Deniau J, Menetrey A, Thierry A (1994) Indirect nucleus accumbens input to the prefrontal cortex via the substantia nigra pars reticulata: a combined anatomical and electrophysiological study in the rat. Neuroscience 61:533-545.
Di Chiara G (2002) Nucleus accumbens shell and core dopamine: differential role in behavior and addiction. Behavioural brain research 137:75-114.
Dommett E, Coizet V, Blaha CD, Martindale J, Lefebvre V, Walton N, Mayhew JE, Overton PG, Redgrave P (2005) How visual stimuli activate dopaminergic neurons at short latency. Science 307:1476-1479.
Dreher JK, Jackson DM (1989) Role of D1 and D2 dopamine receptors in mediating locomotor activity elicited from the nucleus accumbens of rats. Brain research 487:267-277.
Dreyer JK, Herrik KF, Berg RW, Hounsgaard JD (2010) Influence of phasic and tonic dopamine release on receptor activation. J Neurosci 30:14273-14283.
Dugast C, Suaud-Chagny MF, Gonon F (1994) Continuous in vivo monitoring of evoked dopamine release in the rat nucleus accumbens by amperometry. Neuroscience 62:647-654.
Fadda P, Scherma M, Fresu A, Collu M, Fratta W (2005) Dopamine and serotonin release in dorsal striatum and nucleus accumbens is differentially modulated by morphine in DBA/2J and C57BL/6J mice. Synapse 56:29-38.
Fang C, Levinthal D (2009) Near-term liability of exploitation: exploration and exploitation in multistage problems. Organization science 20:538-551.
Faure A, Reynolds SM, Richard JM, Berridge KC (2008) Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. The Journal of neuroscience 28:7184-7192.
Fenno L, Yizhar O, Deisseroth K (2011) The development and application of optogenetics. Annual review of neuroscience 34:389-412.
Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H (2011) A selective role for dopamine in stimulus-reward learning. Nature 469:53-57.
Floresco S, McLaughlin R, Haluk D (2008) Opposing roles for the nucleus accumbens core and shell in cue-induced reinstatement of food-seeking behavior. Neuroscience 154:877-884.
Floresco SB (2013) Prefrontal dopamine and behavioral flexibility: shifting from an “inverted-U” toward a family of functions. Frontiers in Neuroscience 7.
Floresco SB, Magyar O (2006) Mesocortical dopamine modulation of executive functions: beyond working memory. Psychopharmacology 188:567-585.
Floresco SB, Zhang Y, Enomoto T (2009) Neural circuits subserving behavioral flexibility and their relevance to schizophrenia. Behav Brain Res 204:396-409.
Floresco SB, Ghods-Sharifi S, Vexelman C, Magyar O (2006a) Dissociable roles for the nucleus accumbens core and shell in regulating set shifting. J Neurosci 26:2449-2457.
Floresco SB, West AR, Ash B, Moore H, Grace AA (2003) Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. Nature neuroscience 6:968-973.

Floresco SB, Magyar O, Ghods-Sharifi S, Vexelman C, Tse MT (2006b) Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. Neuropsychopharmacology 31:297-309.
Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. Journal of cognitive neuroscience 17:51-72.
Frank MJ, Seeberger LC, O'reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306:1940-1943.
Freeze BS, Kravitz AV, Hammack N, Berke JD, Kreitzer AC (2013) Control of Basal Ganglia Output by Direct and Indirect Pathway Projection Neurons. The Journal of neuroscience 33:18531-18539.
Gallup Jr GG (1965) Aggression in rats as a function of frustrative nonreward in a straight alley. Psychonomic Science.
Gan JO, Walton ME, Phillips PE (2010) Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. Nat Neurosci 13:25-27.
Garris PA, Rebec GV (2002) Modeling fast dopamine neurotransmission in the nucleus accumbens during behavior. Behav Brain Res 137:47-63.
Garris PA, Ciolkowski EL, Pastore P, Wightman R (1994) Efflux of dopamine from the synaptic cleft in the nucleus accumbens of the rat brain. The Journal of neuroscience 14:6084-6093.
George S, Watanabe M, PAOLO TD, Falardeau P, Labrie F, Seeman P (1985) The Functional State of the Dopamine Receptor in the Anterior Pituitary Is in the High Affinity Form*. Endocrinology 117:690-697.
Gerfen CR, Surmeier DJ (2011) Modulation of striatal projection systems by dopamine. Annual review of neuroscience 34:441-466.
Gill TM, Castaneda PJ, Janak PH (2010) Dissociable roles of the medial prefrontal cortex and nucleus accumbens core in goal-directed actions for differential reward magnitude. Cerebral Cortex 20:2884-2899.
Girard I, Garland Jr T (2002) Plasma corticosterone response to acute and chronic voluntary exercise in female house mice. Journal of applied Physiology 92:1553-1561.
Glimcher PW (2011) Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. Proceedings of the National Academy of Sciences 108:15647-15654.
Gonzalez R (2009) Data Analysis for Experimental Design. New York: Guilford Press.
Goto Y, Grace AA (2005) Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. Nat Neurosci 8:805-812.
Goto Y, Otani S, Grace AA (2007) The Yin and Yang of dopamine release: a new perspective. Neuropharmacology 53:583-587.
Grace AA (1991) Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. Neuroscience 41:1-24.
Grace AA, Bunney BS (1984a) The control of firing pattern in nigral dopamine neurons: burst firing. J Neurosci 4:2877-2890.
Grace AA, Bunney BS (1984b) The control of firing pattern in nigral dopamine neurons: single spike firing. J Neurosci 4:2866-2876.

Grace AA, Floresco SB, Goto Y, Lodge DJ (2007) Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. Trends Neurosci 30:220-227.
Haber SN (2003) The primate basal ganglia: parallel and integrative networks. Journal of chemical neuroanatomy 26:317-330.
Haluk DM, Floresco SB (2009) Ventral striatal dopamine modulation of different forms of behavioral flexibility. Neuropsychopharmacology 34:2041-2052.
Hanlon EC, Benca RM, Baldo BA, Kelley AE (2010) REM sleep deprivation produces a motivational deficit for food reward that is reversed by intra-accumbens amphetamine in rats. Brain Res Bull 83:245-254.
Hart AS, Rutledge RB, Glimcher PW, Phillips PE (2014) Phasic Dopamine Release in the Rat Nucleus Accumbens Symmetrically Encodes a Reward Prediction Error Term. The Journal of neuroscience 34:698-704.
Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M (2004) A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. The Journal of neuroscience 24:1660-1665.
Haskell DG (1997) Experiments and a model examining learning in the area-restricted search behavior of ferrets (Mustela putorius furo). Behavioral Ecology 8:448-449.
Heien ML, Phillips PE, Stuber GD, Seipel AT, Wightman RM (2003) Overoxidation of carbon-fiber microelectrodes enhances dopamine adsorption and increases sensitivity. Analyst 128:1413-1419.
Heien MLAV, Johnson MA, Wightman RM (2004) Resolving neurotransmitters detected by fast-scan cyclic voltammetry. Analytical chemistry 76:5697-5704.
Henry DJ, Hu XT, White FJ (1998) Adaptations in the mesoaccumbens dopamine system resulting from repeated administration of dopamine D1 and D2 receptor-selective agonists: relevance to cocaine sensitization. Psychopharmacology 140:233-242.
Hikida T, Kimura K, Wada N, Funabiki K, Nakanishi S (2010) Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. Neuron 66:896-907.
Hills T, Brockie PJ, Maricq AV (2004) Dopamine and glutamate control area-restricted search behavior in Caenorhabditis elegans. The Journal of neuroscience 24:1217-1225.
Hills TT (2006) Animal Foraging and the Evolution of Goal‐Directed Cognition. Cognitive Science 30:3-41.
Hills TT (2011) The Evolutionary Origins of Cognitive Control. Topics in Cognitive Science 3:231-237.
Hills TT, Jones MN, Todd PM (2012) Optimal foraging in semantic memory. Psychological Review 119:431.
Hills TT, Kalff C, Wiener JM (2013) Adaptive Lévy processes and area-restricted search in human foraging. PLoS One 8:e60488.
Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304-309.
Hong S, Hikosaka O (2011) Dopamine-mediated learning and switching in cortico-striatal circuit explain behavioral changes in reinforcement learning. Frontiers in behavioral neuroscience 5.

Hooper KC, Banks DA, Stordahl LJ, White IM, Rebec GV (1997) Quinpirole inhibits striatal and excites pallidal neurons in freely moving rats. Neuroscience Letters 237:69-72.
Horvitz J (2000) Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. Neuroscience 96:651-656.
Hossain MM, Suzuki T, Richardson JR, Kobayashi H (2013) Acute Effects of Pyrethroids on Serotonin Release in the Striatum of Awake Rats: An In Vivo Microdialysis Study. Journal of biochemical and molecular toxicology 27:150-156.
Howland JG, Taepavarapruk P, Phillips AG (2002) Glutamate receptor-dependent modulation of dopamine efflux in the nucleus accumbens by basolateral, but not central, nucleus of the amygdala in rats. The Journal of neuroscience 22:1137-1145.
Humphries MD, Prescott TJ (2010) The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Prog Neurobiol 90:385-417.
Humphries MD, Khamassi M, Gurney K (2012) Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. Frontiers in Neuroscience 6:9.
Hyland B, Reynolds J, Hay J, Perk C, Miller R (2002a) Firing modes of midbrain dopamine cells in the freely moving rat. Neuroscience 114:475-492.
Hyland BI, Reynolds JN, Hay J, Perk CG, Miller R (2002b) Firing modes of midbrain dopamine cells in the freely moving rat. Neuroscience 114:475-492.
Ikemoto S (2007) Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. Brain Res Rev 56:27-78.
Ito R, Dalley JW, Howes SR, Robbins TW, Everitt BJ (2000) Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. J Neurosci 20:7489-7495.
Johansen JP, Wolff SB, Lüthi A, LeDoux JE (2012) Controlling the elements: an optogenetic approach to understanding the neural circuits of fear. Biological Psychiatry 71:1053-1060.
Jonasson Z (2005) Meta-analysis of sex differences in rodent models of learning and memory: a review of behavioral and biological data. Neuroscience & Biobehavioral Reviews 28:811-825.
Jones S, Kornblum JL, Kauer JA (2000) Amphetamine blocks long-term synaptic depression in the ventral tegmental area. The Journal of neuroscience 20:5575-5580.
Joseph J, Wang Y-M, Miles P, Budygin E, Picetti R, Gainetdinov R, Caron M, Wightman R (2002) Dopamine autoreceptor regulation of release and uptake in mouse brain slices in the absence of D< sub> 3</sub> receptors. Neuroscience 112:39-49.
Josiassen RC, Curry LM, Mancall EL (1983) Development of neuropsychological deficits in Huntington's disease. Archives of Neurology 40:791.
Kalivas PW, Duffy P (1991) A comparison of axonal and somatodendritic dopamine release using in vivo dialysis. Journal of neurochemistry 56:961-967.
Kareiva P, Odell G (1987) Swarms of predators exhibit" preytaxis" if individual predators use area-restricted search. American Naturalist:233-270.

Kawagoe KT, Garris PA, Wiedemann DJ, Wightman RM (1992) Regulation of transient dopamine concentration gradients in the microenvironment surrounding nerve terminals in the rat striatum. Neuroscience 51:55-64.
Keck M, Welt T, Müller M, Erhardt A, Ohl F, Toschi N, Holsboer F, Sillaber I (2002) Repetitive transcranial magnetic stimulation increases the release of dopamine in the mesolimbic and mesostriatal system. Neuropharmacology 43:101-109.
Keefe K, Zigmond M, Abercrombie E (1993) In vivo regulation of extracellular dopamine in the neostriatum: influence of impulse activity and local excitatory amino acids. Journal of Neural Transmission/General Section JNT 91:223-240.
Keithley RB, Mark Wightman R, Heien ML (2009) Multivariate concentration determination using principal component regression with residual analysis. TrAC Trends in Analytical Chemistry 28:1127-1136.
Keithley RB, Carelli RM, Wightman RM (2010) Rank estimation and the multivariate analysis of in vivo fast-scan cyclic voltammetric data. Anal Chem 82:5541-5551.
Kelley AE (1999) Functional specificity of ventral striatal compartments in appetitive behaviors. Ann N Y Acad Sci 877:71-90.
Kelley AE, Berridge KC (2002) The neuroscience of natural rewards: relevance to addictive drugs. The Journal of neuroscience 22:3306-3311.
Kemp JM, Powell T (1971) The structure of the caudate nucleus of the cat: light and electron microscopy. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences:383-401.
Kennedy RT (2013) Emerging trends in *in vivo* neurochemical monitoring by microdialysis. Current opinion in chemical biology 17:860-867.
Kerfoot EC, Chattillion EA, Williams CL (2008) Functional interactions between the nucleus tractus solitarius (NTS) and nucleus accumbens shell in modulating memory for arousing experiences. Neurobiology of learning and memory 89:47-60.
Khan AS, Michael AC (2003) Invasive consequences of using micro-electrodes and microdialysis probes in the brain. TrAC Trends in Analytical Chemistry 22:503-508.
Kinsley CH, Madonia L, Gifford GW, Tureski K, Griffin GR, Lowry C, Williams J, Collins J, McLearie H, Lambert KG (1999) Motherhood improves learning and memory: Neural activity in rats is enhanced by pregnancy and the demands of rearing offspring. Nature.
Koeltzow TE, Austin JD, Vezina P (2003) Behavioral sensitization to quinpirole is not associated with increased nucleus accumbens dopamine overflow. Neuropharmacology 44:102-110.
Kravitz AV, Kreitzer AC (2012) Striatal Mechanisms Underlying Movement, Reinforcement, and Punishment. Physiology 27:167-177.
Kravitz AV, Tye LD, Kreitzer AC (2012) Distinct roles for direct and indirect pathway striatal neurons in reinforcement. Nature neuroscience 15:816-U823.
Kravitz AV, Freeze BS, Parker PRL, Kay K, Thwin MT, Deisseroth K, Kreitzer AC (2010) Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature 466:622-626.
Kreitzer A, Berke J (2011) Investigating striatal function through cell-type-specific manipulations. Neuroscience 198:19-26.

Lammel S, Lim BK, Malenka RC (2014) Reward and aversion in a heterogeneous midbrain dopamine system. Neuropharmacology 76:351-359.
Lammel S, Ion DI, Roeper J, Malenka RC (2011) Projection-Specific Modulation of Dopamine Neuron Synapses by Aversive and Rewarding Stimuli. Neuron 70:855-862.
Lammel S, Hetzel A, Hackel O, Jones I, Liss B, Roeper J (2008) Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. Neuron 57:760-773.
Laureiro-Martínez D, Brusoni S, Zollo M (2010) The neuroscientific foundations of the exploration− exploitation dilemma. Journal of Neuroscience, Psychology, and Economics 3:95.
Lecca D, Cacciapaglia F, Valentini V, Gronli J, Spiga S, Di Chiara G (2006) Preferential increase of extracellular dopamine in the rat nucleus accumbens shell as compared to that in the core during acquisition and maintenance of intravenous nicotine self-administration. Psychopharmacology 184:435-446.
Levita L, Dalley JW, Robbins TW (2002) Nucleus accumbens dopamine and learned fear revisited: a review and some new findings. Behavioural brain research 137:115-127.
Liss B, Roeper J (2008) Individual dopamine midbrain neurons: functional diversity and flexibility in health and disease. Brain Research Reviews 58:314-321.
Liu ZH, Shin R, Ikemoto S (2008) Dual role of medial A10 dopamine neurons in affective encoding. Neuropsychopharmacology 33:3010-3020.
Lode T (2000) Functional response and area‐restricted search in a predator: seasonal exploitation of anurans by the European polecat, Mustela putorius. Austral Ecology 25:223-231.
Louilot A, Le Moal M, Simon H (1986) Differential reactivity of dopaminergic neurons in the nucleus accumbens in response to different behavioral situations. An in vivo voltammetric study in free moving rats. Brain research 397:395-400.
Love G, Torrey N, McNamara I, Morgan M, Banks M, Hester NW, Glasper ER, DeVries AC, Kinsley CH, Lambert KG (2005) Maternal experience produces long-lasting behavioral modifications in the rat. Behavioral neuroscience 119:1084.
Lu X-Y, Behnam Ghasemzadeh M, Kalivas P (1997) Expression of D_1 receptor, D_2 receptor, substance P and enkephalin messenger RNAs in the neurons projecting from the nucleus accumbens. Neuroscience 82:767-780.
Lynch WJ, Arizzi MN, Carroll ME (2000) Effects of sex and the estrous cycle on regulation of intravenously self-administered cocaine in rats. Psychopharmacology (Berl) 152:132-139.
Mahler SV, Smith KS, Berridge KC (2007) Endocannabinoid hedonic hotspot for sensory pleasure: Anandamide in nucleus accumbens shell enhances 'liking' of a sweet reward. Neuropsychopharmacology 32:2267-2278.
Maina FK, Mathews TA (2010) Functional fast scan cyclic voltammetry assay to characterize dopamine D2 and D3 autoreceptors in the mouse striatum. ACS chemical neuroscience 1:450-462.
Marcellino D, Kehr J, Agnati LF, Fuxe K (2012) Increased affinity of dopamine for D2‐like versus D1‐like receptors. Relevance for volume transmission in interpreting PET findings. Synapse.

March JG (1991) Exploration and exploitation in organizational learning. Organization science 2:71-87.
Margolis EB, Hjelmstad GO, Bonci A, Fields HL (2003) Kappa-opioid agonists directly inhibit midbrain dopaminergic neurons. J Neurosci 23:9981-9986.
Margolis EB, Lock H, Hjelmstad GO, Fields HL (2006a) The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? J Physiol 577:907-924.
Margolis EB, Lock H, Chefer VI, Shippenberg TS, Hjelmstad GO, Fields HL (2006b) Kappa opioids selectively control dopaminergic neurons projecting to the prefrontal cortex. Proc Natl Acad Sci U S A 103:2938-2942.
Matamales M, Bertran-Gonzalez J, Salomon L, Degos B, Deniau J-M, Valjent E, Hervé D, Girault J-A (2009) Striatal medium-sized spiny neurons: identification by nuclear staining and study of neuronal subpopulations in BAC transgenic mice. PLoS One 4:e4770.
Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459:837-841.
Matsumoto M, Takada M (2013) Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. Neuron 79:1011-1024.
Maurice N, Deniau J, Menetrey A, Glowinski J, Thierry A (1997) Position of the ventral pallidum in the rat prefrontal cortex–basal ganglia circuit. Neuroscience 80:523-534.
McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. Neuron 38:339-346.
McCullough L, Sokolowski J, Salamone J (1993) A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. Neuroscience 52:919-925.
McCutcheon JE, Ebner SR, Loriaux AL, Roitman MF (2012) Encoding of Aversion by Dopamine and the Nucleus Accumbens. Frontiers in Neuroscience 6:137.
McGregor A, Roberts D (1993) Dopaminergic antagonism within the nucleus accumbens or the amygdala produces differential effects on intravenous cocaine self-administration under fixed and progressive ratio schedules of reinforcement. Brain research 624:245-252.
Meck WH, Benson AM (2002) Dissecting the brain's internal clock: how frontal–striatal circuitry keeps time and shifts attention. Brain and cognition 48:195-211.
Meredith GE, Baldo BA, Andrezjewski ME, Kelley AE (2008) The structural basis for mapping behavior onto the ventral striatum and its subdivisions. Brain Struct Funct 213:17-27.
Michael D, Travis ER, Wightman RM (1998) Peer Reviewed: Color Images for Fast-Scan CV Measurements in Biological Systems. Analytical chemistry 70:586A-592A.
Miller NE, Stevenson SS (1936) Agitated behavior of rats during experimental extinction and a curve of spontaneous recovery. Journal of Comparative Psychology 21:205.
Montoya A, Price BH, Menear M, Lepage M (2006) Brain imaging and cognitive dysfunctions in Huntington's disease. Journal of Psychiatry and Neuroscience 31:21.

Moreno M, Economidou D, Mar AC, Lopez-Granero C, Caprioli D, Theobald DE, Fernando A, Newman AH, Robbins TW, Dalley JW (2013) Divergent effects of D(2/3) receptor activation in the nucleus accumbens core and shell on impulsivity and locomotor activity in high and low impulsive rats. Psychopharmacology (Berl).

Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H (2006) Midbrain dopamine neurons encode decisions for future action. Nat Neurosci 9:1057-1063.

Mustard JA, Beggs KT, Mercer AR (2005) Molecular biology of the invertebrate dopamine receptors. Archives of insect biochemistry and physiology 59:103-117.

Nandi P, Lunte SM (2009) Recent trends in microdialysis sampling integrated with conventional and microanalytical systems for monitoring biological events: a review. Analytica chimica acta 651:1-14.

Natesan S, Reckless GE, Nobrega JN, Fletcher PJ, Kapur S (2006) Dissociation between in vivo occupancy and functional antagonism of dopamine D2 receptors: comparing aripiprazole to other antipsychotics in animal models. Neuropsychopharmacology 31:1854-1863.

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

Nishino H, Ono T, Muramoto K, Fukuda M, Sasaki K (1987) Neuronal activity in the ventral tegmental area (VTA) during motivated bar press feeding in the monkey. Brain research 413:302-313.

Niv Y (2007) Cost, Benefit, Tonic, Phasic. What do response rates tell us about dopamine and motivation? Annals of the New York Academy of Sciences 1104:357-376.

Niv Y, Schoenbaum G (2008) Dialogues on prediction errors. Trends in cognitive sciences 12:265-272.

Niv Y, Daw ND, Dayan P (2005) How fast to work: Response vigor, motivation and tonic dopamine. In: NIPS, pp 1019-1026.

Niv Y, Daw ND, Dayan P (2006) Choice values. Nat Neurosci 9:987-988.

Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology (Berl) 191:507-520.

O'Connor JJ, Lowry JP (2012) A comparison of the effects of the dopamine partial agonists aripiprazole and (−)-3-PPP with quinpirole on stimulated dopamine release in the rat striatum: studies using fast cyclic voltammetry< i> in vitro</i>. European Journal of Pharmacology.

Oleson EB, Gentry RN, Chioma VC, Cheer JF (2012) Subsecond Dopamine Release in the Nucleus Accumbens Predicts Conditioned Punishment and Its Successful Avoidance. The Journal of neuroscience 32:14804-14808.

Overton P, Vautrelle N, Redgrave P (2014) Sensory regulation of dopaminergic cell activity: phenomenology, circuitry and function. Neuroscience.

Owesson-White CA, Roitman MF, Sombers LA, Belle AM, Keithley RB, Peele JL, Carelli RM, Wightman RM (2012) Sources contributing to the average extracellular concentration of dopamine in the nucleus accumbens. Journal of neurochemistry.

Palmiter RD (2007) Is dopamine a physiologically relevant mediator of feeding behavior? Trends in neurosciences 30:375-381.

Pan WX, Schmidt R, Wickens JR, Hyland BI (2005) Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. J Neurosci 25:6235-6242.
Pan WX, Schmidt R, Wickens JR, Hyland BI (2008) Tripartite mechanism of extinction suggested by dopamine neuron activity and temporal difference model. The Journal of neuroscience 28:9619-9631.
Papini MR, Dudley RT (1997) Consequences of surprising reward omissions. Review of General Psychology 1:175.
Paulson P, Camp D, Robinson T (1991) Time course of transient behavioral depression and persistent behavioral sensitization in relation to regional brain monoamine concentrations during amphetamine withdrawal in rats. Psychopharmacology 103:480-492.
Pawluski JL, Walker SK, Galea LAM (2006) Reproductive experience differentially affects spatial reference and working memory performance in the mother. Hormones and behavior 49:143-149.
Paxinos G, Watson C (1998) The rat brain in stereotaxic coordinates. San Diego: Academic Press.
Peters JL, Miner LH, Michael AC, Sesack SR (2004) Ultrastructure at carbon fiber microelectrode implantation sites after acute voltammetric measurements in the striatum of anesthetized rats. J Neurosci Methods 137:9-23.
Pezze MA, Dalley JW, Robbins TW (2007) Differential roles of dopamine D1 and D2 receptors in the nucleus accumbens in attentional performance on the five-choice serial reaction time task. Neuropsychopharmacology 32:273-283.
Phillips GD, Howes SR, Whitelaw RB, Everitt BJ, Robbins T (1995) Analysis of the effects of intra-accumbens SKF-38393 and LY-171555 upon the behavioural satiety sequence. Psychopharmacology 117:82-90.
Phillips PE, Walton ME, Jhou TC (2007a) Calculating utility: preclinical evidence for cost–benefit analysis by mesolimbic dopamine. Psychopharmacology 191:483-495.
Phillips PE, Walton ME, Jhou TC (2007b) Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. Psychopharmacology (Berl) 191:483-495.
Phillips PE, Robinson DL, Stuber GD, Carelli RM, Wightman RM (2003) Real-time measurements of phasic changes in extracellular dopamine concentration in freely moving rats by fast-scan cyclic voltammetry. Methods Mol Med 79:443-464.
Pierce RC, Kalivas PW (1995) Amphetamine produces sensitized increases in locomotion and extracellular dopamine preferentially in the nucleus accumbens shell of rats administered repeated cocaine. Journal of Pharmacology and Experimental Therapeutics 275:1019-1029.
Pierce RC, Duffy P, Kalivas PW (1995) Sensitization to cocaine and dopamine autoreceptor subsensitivity in the nucleus accumbens. Synapse 20:33-36.
Porter-Stransky KA, Seiler JL, Day JJ, Aragona BJ (2013) Development of behavioral preferences for the optimal choice following unexpected reward omission is mediated by a reduction of D2‐like receptor tone in the nucleus accumbens. European Journal of Neuroscience.

Porter-Stransky KA, Wescott SA, Hershman M, Badrinarayan A, Vander Weele CM, Lovic V, Aragona BJ (2011) Cocaine must enter the brain to evoke unconditioned dopamine release within the nucleus accumbens shell. Neurosci Lett 504:13-17.
Prinssen EP, Kleven MS, Koek W (1996) Effects of dopamine antagonists in a two-way active avoidance procedure in rats: interactions with 8-OH-DPAT, ritanserin, and prazosin. Psychopharmacology 128:191-197.
Pulvirenti L, Berrier R, Kreifeldt M, Koob GK (1994) Modulation of locomotor activity by NMDA receptors in the nucleus accumbens core and shell regions of the rat. Brain research 664:231-236.
Purgert RJ, Wheeler DS, McDannald MA, Holland PC (2012) Role of Amygdala Central Nucleus in Aversive Learning Produced by Shock or by Unexpected Omission of Food. The Journal of neuroscience 32:2461-2472.
Pyke G (1984) Optimal Foraging Theory: A Critical Review. Annual Review of Ecology and Systematics 15:523-575.
Rada PV, Mark GP, Hoebel BG (1998) Dopamine release in the nucleus accumbens by hypothalamic stimulation-escape behavior. Brain research 782:228-234.
Ragozzino ME (2002) The effects of dopamine D(1) receptor blockade in the prelimbic-infralimbic areas on behavioral flexibility. Learn Mem 9:18-28.
Ragozzino ME (2007) The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. Annals of the New York Academy of Sciences 1121:355-375.
Ragozzino ME, Kim J, Hassert D, Minniti N, Kiang C (2003) The contribution of the rat prelimbic-infralimbic areas to different forms of task switching. Behavioral neuroscience 117:1054.
Ramot A, Akirav I (2012) Cannabinoid receptors activation and glucocorticoid receptors deactivation in the amygdala prevent the stress-induced enhancement of a negative learning experience. Neurobiology of learning and memory.
Rana AQ, Masroor MS, Khan AS (2013) A review of methods used to study cognitive deficits in Parkinson's disease. Neurol Res 35:1-6.
Ranaldi R, Beninger RJ (1994) The effects of systemic and intracerebral injections of D1 and D2 agonists on brain stimulation reward. Brain Res 651:283-292.
Ranaldi R, Pocock D, Zereik R, Wise RA (1999) Dopamine fluctuations in the nucleus accumbens during maintenance, extinction, and reinstatement of intravenous D-amphetamine self-administration. The Journal of neuroscience 19:4102-4109.
Reading PJ, Dunnett SB (1991) The effects of excitotoxic lesions of the nucleus accumbens on a matching to position task. Behav Brain Res 46:17-29.
Redgrave P, Gurney K (2006) The short-latency dopamine signal: a role in discovering novel actions? Nature Reviews Neuroscience 7:967-975.
Redgrave P, Gurney K, Reynolds J (2008) What is reinforced by phasic dopamine signals? Brain Research Reviews 58:322-339.
Reid AK, Staddon J (1998) A dynamic route finder for the cognitive map. Psychological Review 105:585.
Reiner A, Medina L, Veenman CL (1998) Structural and functional evolution of the basal ganglia in vertebrates. Brain Research Reviews 28:235-285.

Reynolds SM, Berridge KC (2008) Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. Nature neuroscience 11:423-425.
Richard JM, Berridge KC (2011) Nucleus Accumbens Dopamine/Glutamate Interaction Switches Modes to Generate Desire versus Dread: D1 Alone for Appetitive Eating But D1 and D2 Together for Fear. The Journal of neuroscience 31:12866-12879.
Richfield EK, Penney JB, Young AB (1989) Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. Neuroscience 30:767-777.
Robertson G, Jian M (1995) D_1 and D_2 dopamine receptors differentially increase fos-like immunoreactivity in accumbal projections to the ventral pallidum and midbrain. Neuroscience 64:1019-1034.
Robinson DL, Wightman RM (2007) Rapid Dopamine Release in Freely Moving Rats.
Robinson DL, Venton BJ, Heien MLAV, Wightman RM (2003) Detecting subsecond dopamine release with fast-scan cyclic voltammetry in vivo. Clinical chemistry 49:1763.
Robinson S, Sandstrom SM, Denenberg VH, Palmiter RD (2005) Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. Behavioral neuroscience 119:5.
Robinson TE, Camp DM (1987) Long-lasting effects of escalating doses of *d*-amphetamine on brain monoamines, amphetamine-induced stereotyped behavior and spontaneous nocturnal locomotion. Pharmacology Biochemistry and Behavior 26:821-827.
Robinson TE, Berridge KC (1993) The neural basis of drug craving: an incentive-sensitization theory of addiction. Brain Res Brain Res Rev 18:247-291.
Robinson TE, Yager LM, Cogan ES, Saunders BT (2014) On the motivational properties of reward cues: Individual differences. Neuropharmacology 76:450-459.
Roeper J (2013) Dissecting the diversity of midbrain dopamine neurons. Trends in neurosciences 36:336-342.
Roesch MR, Calu DJ, Schoenbaum G (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat Neurosci 10:1615-1624.
Salamone JD, Correa M (2012) The mysterious motivational functions of mesolimbic dopamine. Neuron 76:470-485.
Salamone JD, Cousins MS, Bucher S (1994) Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. Behavioural brain research 65:221-229.
Salamone JD, Correa M, Farrar A, Mingote SM (2007) Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. Psychopharmacology (Berl) 191:461-482.
Salinas JA, White NM (1998) Contributions of the hippocampus, amygdala, and dorsal striatum to the response elicited by reward reduction. Behavioral Neuroscience; Behavioral Neuroscience 112:812.

Salinas JA, Gold PE (2005) Glucose regulation of memory for reward reduction in young and aged rats. Neurobiol Aging 26:45-52.
Salinas JA, Packard MG, McGaugh JL (1993) Amygdala modulates memory for changes in reward magnitude - reversible post-training inactivation with lidocaine attenuates the response to a reduction in reward. Behavioural brain research 59:153-159.
Salinas JA, Williams CL, McGaugh JL (1996) Peripheral Post-training Administration of 4-OH Amphetamine Enhances Retention of a Reduction in Reward Magnitude. Neurobiology of learning and memory 65:192-195.
Salinas JA, Introini-Collison IB, Dalmaz C, McGaugh JL (1997) Posttraining intraamygdala infusions of oxotremorine and propranolol modulate storage of memory for reductions in reward magnitude. Neurobiology of learning and memory 68:51-59.
Sam PM, Justice JB, Jr. (1996) Effect of general microdialysis-induced depletion on extracellular dopamine. Anal Chem 68:724-728.
Sarter M, Fritschy JM (2008) Reporting statistical methods and statistical results in EJN. European Journal of Neuroscience 28:2363-2364.
Sastre A, Reilly S (2006) Excitotoxic lesions of the gustatory thalamus eliminate consummatory but not instrumental successive negative contrast in rats. Behavioural brain research 170:34-40.
Saunders BT, Robinson TE (2012) The role of dopamine in the accumbens core in the expression of Pavlovian‐conditioned responses. European Journal of Neuroscience 36:2521-2532.
Schall JD (2003) Neural correlates of decision processes: neural and mental chronometry. Current opinion in neurobiology 13:182-186.
Schultz KN, Kennedy RT (2008) Time-resolved microdialysis for in vivo neurochemical measurements and other applications. Annu Rev Anal Chem 1:627-661.
Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1-27.
Schultz W (2002) Getting formal with dopamine and reward. Neuron 36:241-263.
Schultz W (2007) Multiple dopamine functions at different time courses. Annu Rev Neurosci 30:259-288.
Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593-1599.
Seeman P (2009) Schizophrenia model of elevated $D2^{High}$ receptors: Haloperidol reverses the amphetamine-induced elevation in dopamine $D2^{High}$. Schizophrenia research 109:191-192.
Seeman P (2011) All roads to schizophrenia lead to dopamine supersensitivity and elevated dopamine D2High receptors. CNS neuroscience & therapeutics 17:118-132.
Seeman P, Weinshenker D, Quirion R, Srivastava LK, Bhardwaj SK, Grandy DK, Premont RT, Sotnikova TD, Boksa P, El-Ghundi M (2005) Dopamine supersensitivity correlates with D2High states, implying many paths to psychosis. Proceedings of the National Academy of Sciences of the United States of America 102:3513-3518.

Shou M, Ferrario CR, Schultz KN, Robinson TE, Kennedy RT (2006) Monitoring dopamine in vivo by microdialysis sampling and on-line CE-laser-induced fluorescence. Analytical chemistry 78:6717-6725.
Sinkala E, McCutcheon JE, Schuck MJ, Schmidt E, Roitman MF, Eddington DT (2012) Electrode calibration with a microfluidic flow cell for fast-scan cyclic voltammetry. Lab on a Chip 12:2403-2408.
Skelin I, Hakstol R, VanOyen J, Mudiayi D, Molina LA, Holec V, Hong NS, Euston DR, McDonald RJ, Gruber AJ (2014) Lesions of dorsal striatum eliminate lose‐switch responding but not mixed‐response strategies in rats. European Journal of Neuroscience.
Skinner BF (1953) Science and human behavior: Free Press.
Slaney TR, Mabrouk OS, Porter-Stransky KA, Aragona BJ, Kennedy RT (2013) Chemical Gradients within Brain Extracellular Space Measured using Low Flow Push–Pull Perfusion Sampling in Vivo. ACS chemical neuroscience.
Smeets WJ, González A (2000) Catecholamine systems in the brain of vertebrates: new perspectives through a comparative approach. Brain Research Reviews 33:308-379.
Smith KS, Berridge KC, Aldridge JW (2011) Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. Proceedings of the National Academy of Sciences 108:E255-E264.
Song P, Mabrouk OS, Hershey ND, Kennedy RT (2012) In Vivo Neurochemical Monitoring using Benzoyl Chloride Derivatization and Liquid Chromatography–Mass Spectrometry. Analytical chemistry.
Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013) A causal link between prediction errors, dopamine neurons and learning. Nature neuroscience.
Stephens DW, Krebs JR (1986) Foraging theory: Princeton Univ Pr.
Stopper CM, Khayambashi S, Floresco SB (2013) Receptor-Specific Modulation of Risk-Based Decision Making by Nucleus Accumbens Dopamine. Neuropsychopharmacology.
Stout D (2010) The evolution of cognitive control. Topics in Cognitive Science 2:614-630.
Stratmann JA, Craft RM (1997) Intracranial self-stimulation in female and male rats: no sex differences using a rate-independent procedure. Drug and Alcohol Dependence 46:31-40.
Stuber GD, Wightman RM, Carelli RM (2005) Extinction of cocaine self-administration reveals functionally and temporally distinct dopaminergic signals in the nucleus accumbens. Neuron 46:661-669.
Sugam JA, Saddoris MP, Carelli RM (2013) Nucleus Accumbens Neurons Track Behavioral Preferences and Reward Outcomes During Risky Decision Making. Biological Psychiatry.
Sugam JA, Day JJ, Wightman RM, Carelli RM (2012) Phasic Nucleus Accumbens Dopamine Encodes Risk-Based Decision-Making Behavior. Biological Psychiatry.
Sulzer D (2011) How addictive drugs disrupt presynaptic dopamine neurotransmission. Neuron 69:628-649.

Sulzer D, Pothos EN (2000) Regulation of quantal size by presynaptic mechanisms. Reviews in the Neurosciences 11:159-212.
Surmeier D, Carrillo-Reid L, Bargas J (2011) Dopaminergic modulation of striatal neurons, circuits, and assemblies. Neuroscience 198:3-18.
Surmeier DJ, Ding J, Day M, Wang Z, Shen W (2007) D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. Trends in neurosciences 30:228-235.
Sutcliffe JS (2011) Female Rats Are Smarter than Males: Influence of Test, Oestrogen Receptor Subtypes and Glutamate. Biological Basis of Sex Differences in Psychopharmacology:37-56.
Swanson CJ, Heath S, Stratford TR, Kelley AE (1997) Differential behavioral responses to dopaminergic stimulation of nucleus accumbens subregions in the rat. Pharmacol Biochem Behav 58:933-945.
Taghzouti K, Louilot A, Herman JP, Le Moal M, Simon H (1985) Alternation behavior, spatial discrimination, and reversal disturbances following 6-hydroxydopamine lesions in the nucleus accumbens of the rat. Behav Neural Biol 44:354-363.
Tepper JM, Bolam JP (2004) Functional diversity and specificity of neostriatal interneurons. Current opinion in neurobiology 14:685-692.
Thoma P, Wiebel B, Daum I (2007) Response inhibition and cognitive flexibility in schizophrenia with and without comorbid substance use disorder. Schizophrenia research 92:168-180.
Tinklepaugh OL (1928) An experimental study of representative factors in monkeys. Journal of Comparative Psychology 8:197.
Tropp J, Markus EJ (2001) Effects of mild food deprivation on the estrous cycle of rats. Physiology & Behavior 73:553-559.
Ungless MA, Grace AA (2012) Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. Trends in neurosciences 35:422-430.
Ungless MA, Magill PJ, Bolam JP (2004) Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science 303:2040-2042.
Van Haaren F, Wouters M, Van de Poll NE (1987) Absence of behavioral differences between male and female rats in different radial-maze procedures. Physiology & Behavior 39:409-412.
Van Haaren F, Van Hest A, Heinsbroek RPW (1990) Behavioral differences between male and female rats: effects of gonadal hormones on learning and memory. Neuroscience & Biobehavioral Reviews 14:23-33.
Van Hest A, Van Haaren F, Van de Poll NE (1988) The behavior of male and female Wistar rats pressing a lever for food is not affected by sex differences in food motivation. Behavioural brain research 27:215-221.
Van Hest A, Van Haaren F, Van de Poll NE (1989) Operant conditioning of response variability in male and female Wistar rats. Physiology & Behavior 45:551-555.
van Wieringen J-P, Booij J, Shalgunov V, Elsinga P, Michel MC (2013) Agonist high- and low-affinity states of dopamine D2 receptors: methods of detection and clinical implications. Naunyn-Schmiedeberg's archives of pharmacology 386:135-154.

Vander Weele CM, Porter-Stransky KA, Mabrouk OS, Lovic V, Singer BF, Kennedy RT, Aragona BJ (submitted) Differential responses in rapid dopamine transmission within the nucleus accumbens following morphine and oxycodone delivery.
Veeneman M, van Ast M, Broekhoven M, Limpens J, Vanderschuren L (2012) Seeking–taking chain schedules of cocaine and sucrose self-administration: effects of reward size, reward omission, and α-flupenthixol. Psychopharmacology 220:771-785.
Vezina P (2004) Sensitization of midbrain dopamine neuron reactivity and the self-administration of psychomotor stimulant drugs. Neuroscience & Biobehavioral Reviews 27:827-839.
Vindas MA, Folkedal O, Kristiansen TS, Stien LH, Braastad BO, Mayer I, Overli O (2012) Omission of expected reward agitates Atlantic salmon (Salmo salar). Anim Cogn 15:903-911.
Voon V, Pessiglione M, Brezing C, Gallea C, Fernandez HH, Dolan RJ, Hallett M (2010) Mechanisms underlying dopamine-mediated reward bias in compulsive behaviors. Neuron 65:135-142.
Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. Nature 412:43-48.
Wanat MJ, Kuhnen CM, Phillips PE (2010) Delays conferred by escalating costs modulate dopamine release to rewards but not their predictors. J Neurosci 30:12020-12027.
Wanat MJ, Bonci A, Phillips PE (2013) CRF acts in the midbrain to attenuate accumbens dopamine release to rewards but not their predictors. Nature neuroscience 16:383-385.
Wanat MJ, Hopf FW, Stuber GD, Phillips PE, Bonci A (2008) Corticotropin-releasing factor increases mouse ventral tegmental area dopamine neuron firing through a protein kinase C-dependent enhancement of Ih. J Physiol 586:2157-2170.
Wang M, Hershey ND, Mabrouk OS, Kennedy RT (2011) Collection, storage, and electrophoretic analysis of nanoliter microdialysis samples collected from awake animals in vivo. Analytical and bioanalytical chemistry 400:2013-2023.
Wang M, Roman GT, Schultz K, Jennings C, Kennedy RT (2008) Improved temporal resolution for in vivo microdialysis by using segmented flow. Analytical chemistry 80:5607-5615.
Weimerskirch H, Pinaud D, Pawlowski F, Bost CA (2007) Does Prey Capture Induce Area‐Restricted Search? A Fine‐Scale Study Using GPS in a Marine Predator, the Wandering Albatross. The American Naturalist 170:734-743.
Wheeler RA, Aragona BJ, Fuhrmann KA, Jones JL, Day JJ, Cacciapaglia F, Wightman RM, Carelli RM (2011) Cocaine Cues Drive Opposing Context-Dependent Shifts in Reward Processing and Emotional State. Biological Psychiatry.
White J, Tobin TR, Bell WJ (1984) Local search in the housefly *Musca domestica* after feeding on sucrose. Journal of Insect Physiology 30:477-487.
Wickens JR, Horvitz JC, Costa RM, Killcross S (2007) Dopaminergic mechanisms in actions and habits. The Journal of neuroscience 27:8181-8183.
Winter S, Dieckmann M, Schwabe K (2009) Dopamine in the prefrontal cortex regulates rats behavioral flexibility to changing reward value. Behavioural brain research 198:206-213.

Wise RA (2004) Dopamine, learning and motivation. Nat Rev Neurosci 5:483-494.
Wise RA (2008) Dopamine and reward: the anhedonia hypothesis 30 years on. Neurotoxicity research 14:169-183.
Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C (2011) Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. Neuron 72:721-733.
Wolfe JM (2013) When is it time to move to the next raspberry bush? Foraging rules in human visual search. Journal of vision 13.
Wolterink G, Phillips G, Cador M, Donselaar-Wolterink I, Robbins TW, Everitt BJ (1993) Relative roles of ventral striatal D1 and D2 dopamine receptors in responding with conditioned reinforcement. Psychopharmacology (Berl) 110:355-364.
Wyvell CL, Berridge KC (2000) Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: enhancement of reward "wanting" without enhanced "liking" or response reinforcement. The Journal of neuroscience 20:8122-8130.
Yager LM, Robinson TE (2013) A classically conditioned cocaine cue acquires greater control over motivated behavior in rats prone to attribute incentive salience to a food cue. Psychopharmacology 226:217-228.
Young AM (2004) Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: studies using 1 min microdialysis in rats. Journal of neuroscience methods 138:57-63.
Young EJ, Williams CL (2010) Valence Dependent Asymmetric Release of Norepinephrine in the Basolateral Amygdala. Behavioral neuroscience 124:633-644.
Zahm DS (1999) Functional-anatomical Implications of the Nucleus Accumbens Core and Shell Subterritories. Annals of the New York Academy of Sciences 877:113-128.
Zhang L, Doyon WM, Clark JJ, Phillips PE, Dani JA (2009) Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum. Mol Pharmacol 76:396-404.
Zhang M, Balmadrid C, Kelley AE (2003) Nucleus accumbens opioid, GABaergic, and dopaminergic modulation of palatable food motivation: contrasting effects revealed by a progressive ratio study in the rat. Behavioral neuroscience 117:202.
Zhang X, Rauch A, Lee H, Xiao H, Rainer G, Logothetis NK (2007) Capillary hydrophilic interaction chromatography/mass spectrometry for simultaneous determination of multiple neurotransmitters in primate cerebral cortex. Rapid Communications in Mass Spectrometry 21:3621-3628.
Zhou L, Furuta T, Kaneko T (2003) Chemical organization of projection neurons in the rat accumbens nucleus and olfactory tubercle. Neuroscience 120:783-798.