PHASIC NUCLEUS ACCUMBENS DOPAMINE ENCODES
RISK-BASED DECISION MAKING BEHAVIOR

Jonathan Adam Sugam

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology.

Chapel Hill
2011

Approved By:

Advisor: Dr. Regina M. Carelli

Committee member: Dr. Todd E. Thiele

Committee member: Dr. Donald T. Lysle

©2011
Jonathan Adam Sugam
ALL RIGHTS RESERVED

ABSTRACT

JONATHAN SUGAM: Phasic nucleus accumbens dopamine encodes
risk-based decision making behavior
(Under the direction of Regina M. Carelli)

When deciding to act organisms must correctly evaluate the costs and benefits of action selection in order to optimize resources. Several factors influence the evaluation of reward value including the behavioral cost, the delay to, probability of, and magnitude of reinforcement. Dopamine (DA) transmission within the nucleus accumbens (NAc) has been implicated in reward learning and value-based decision making. The mesolimbic DA system has been shown to encode information about explicit value associated with behavioral performance that biases responses towards the most valuable options available; however it is unclear whether subjective value associated with risk-taking behavior is encoded by DA release. Here, phasic DA release in the NAc was monitored using fast scan cyclic voltammetry during a risk-based decision making task in which visual cues predicted the opportunity to respond for small certain or larger uncertain rewards. Reward predictive cues evoked increases in DA concentration in the NAc core that scaled with each animal's preferred reward contingency. Further, DA signaling in the NAc core tracked the unexpected presentation or omission of rewards following prediction error theory. Similar processing was not observed in the NAc shell. These results suggest that the dopaminergic projections to the NAc core encode the subjective value of future rewards that may function to optimize choices during decision making and risk-taking behavior.

ACKNOWLEDGEMENTS

I would like to thank Jeremy J. Day and Michael P. Saddoris for their technical expertise and guidance and Katherine Fuhrmann and Laura Ciompi for technical assistance.

This research was supported by NIDA DA017318 to RMC.

TABLE OF CONTENTS

LIST OF FIGURES.....	vii
----------------------	-----

Chapter

I. INTRODUCTION.....	1
The nucleus accumbens and goal-directed behavior.....	1
Associative learning properties of nucleus accumbens activity and dopamine signaling.....	2
Modeling value-based decision making.....	4
Value-based decision making and the mesolimbic dopamine system.....	6
II. METHODS.....	8
Animals.....	8
Behavioral training.....	8
Surgery.....	11
Fast-scan cyclic voltammetry.....	11
Signal identification and separation.....	12
Data analysis.....	13
Histology.....	16
III. RESULTS.....	17
Behavioral performance on risky decision making task.....	17
Reward-associated discriminative stimuli evoke	

phasic dopamine signals in the NAc.....	18
Cue-evoked dopamine signals within the nucleus accumbens core reflect reward preferences.....	19
Dopamine signaling in the NAc core tracks prediction errors during risky decision making.....	21
IV. DISCUSSION.....	23
REFERENCES.....	39

LIST OF FIGURES

Figure

1. Risky decision making task.....	31
2. Behavioral performance during risky decision making task.....	32
3. Phasic dopamine signaling for a risk-averse rat.....	33
4. Phasic dopamine signaling for a risk-prone rat.....	34
5. Correlation of risk-taking behavior and cue-evoked dopamine signaling.....	35
6. Phasic dopamine signaling for preferred versus nonpreferred cues and reward presentation in the nucleus accumbens core.....	36
7. Anatomical distribution of recording sites.....	37
8. Behavioral performance and phasic dopamine signaling during risky decision making task for nucleus accumbens shell group.....	38

ABBREVIATIONS

ANOVA	Analysis of Variance
BLA	Basolateral Amygdala
DA	Dopamine
FR	Fixed Ratio
NAc	Nucleus Accumbens
OFC	Orbitofrontal Cortex
PFC	Prefrontal Cortex
SEM	Standard Error of the Mean
VTA	Ventral Tegmental Area

CHAPTER 1

INTRODUCTION

The ability to learn associations between cues and positive outcomes allows an organism to locate and secure resources for survival such as food, shelter, and mates. In order to appropriately seek and secure necessary resources, organisms must learn to evaluate the costs and benefits of performing a behavior to determine the best option to maximize their behavioral output (Green and Myerson 2004; Cardinal 2006; Phillips, Walton et al. 2007; Rangel, Camerer et al. 2008). When making these decisions the expected value of performing a task is a key component that biases responses towards the more valuable option available (Phillips, Walton et al. 2007; Rangel, Camerer et al. 2008). Reward value is thus a critical factor that mediates normal decision making behaviors while aberrations in reward valuation processes result in maladaptive behaviors such as drug addiction. For example, drug addiction has been described as an aberrant learning process that is associated with risk-taking behavior and poor decision making (Jentsch and Taylor 1999; Robbins and Everitt 1999). Thus, understanding value-based decision making could help further understand a variety of normal and abnormal human activities, including drug taking, food seeking, social attachment, and sexual behavior.

The nucleus accumbens and goal-directed behavior

The ability to form and use the associations between cues and positive outcomes to drive goal-directed behavior requires brain circuitry that can integrate sensory information

and motivational information to mediate motor output. The nucleus accumbens (NAc) is a structure that is uniquely situated to allow for this processing and is implicated in the mediation of goal-directed behaviors (Kelley 2004). The NAc is part a larger corticolimbic reward circuit that receives glutamatergic input from the hippocampus, prefrontal cortex (PFC), basolateral amygdala (BLA) and a dense dopaminergic input from the ventral tegmental area (VTA). These structures and associated projections to the NAc provide reward related information which functions to influence goal-directed behaviors. The NAc integrates this information and impacts behavior through its connections with motor-related regions (e.g., ventral pallidum). This connectivity supports the role of the NAc as a 'limbic motor interface' that enables the execution of motivated behaviors (Mogenson, Jones et al. 1980). In support of its role in the mediation of goal-directed behaviors, *in vivo* electrophysiological and electrochemical techniques have demonstrated that cellular activity (Carelli, King et al. 1993; Carelli and Deadwyler 1994; Carelli 2000; Nicola, Yun et al. 2004; Roitman, Wheeler et al. 2005) and phasic dopamine (DA) release (Phillips, Stuber et al. 2003; Roitman, Stuber et al. 2004) in the NAc are activated during behavioral events associated with obtaining and consuming food and drug rewards. Further, pharmacological and lesion studies have demonstrated that the NAc plays a critical role in both the appetitive and consummatory phases of goal-directed actions (Berridge and Robinson 1998; Stratford and Kelley 1999; Kelley 2004; Nicola, Taha et al. 2005), suggesting a critical role for the NAc in motivated behaviors.

Associative learning properties of nucleus accumbens activity and dopamine signaling

The NAc and its dopaminergic input from the VTA have been shown to be critical for learning association between cues and positive outcomes which promotes goal-directed

behavior (Berridge and Robinson 1998; Nicola, Yun et al. 2004; Nicola, Taha et al. 2005; Day, Wheeler et al. 2006; Day, Roitman et al. 2007; Hollander and Carelli 2007). Mesolimbic DA neurons show increased phasic activity to reward paired cues and encode violations in prediction following reward paired cues (Schultz, Dayan et al. 1997; Schultz 1998; Waelti, Dickinson et al. 2001). These DA neurons project to the NAc, as it has been shown that phasic DA release in the NAc encodes associations between cues and outcomes (Day, Roitman et al. 2007; Stuber, Klanker et al. 2008), and is essential for stimulus-outcome learning (Sombers, Beyene et al. 2009; Zellner, Kest et al. 2009). As the mesolimbic DA system is a neuromodulator of NAc activity (Yun, Wakabayashi et al. 2004; Goto and Grace 2005), NAc neurons also display increased and/or decreased cell firing to cues that predict future rewards (Carelli 2000; Carelli 2004; Nicola, Yun et al. 2004; Roitman, Wheeler et al. 2005; Day, Wheeler et al. 2006; Jones, Day et al. 2010).

The mesolimbic DA circuitry and associated NAc neural responses are critical in encoding specific characteristics of reward associations that biases future actions. For example, NAc neurons exhibit differential activity for drug versus natural rewards (Carelli, Ijames et al. 2000) and for appetitive versus aversive outcomes (Roitman, Wheeler et al. 2005). Further, ventral striatal neurons display differential activity to cue presentation based on the value of future outcomes (Schultz, Apicella et al. 1992; Cromwell and Schultz 2003; Kim, Sul et al. 2009). Mesolimbic DA neurons also display increased activity for cues that predict rewards based on magnitude, delay (Roesch, Calu et al. 2007), probability (Fiorillo, Tobler et al. 2003), and expected value (Tobler, Fiorillo et al. 2005), implicating a critical role for this system in encoding reward value.

Modeling value-based decision making

Rarely in the environment do organisms encounter situations in which simple stimulus-outcome associations are in effect, and thus organisms must learn to evaluate the risks and benefits of action performance when making decisions to maximize their behavioral output (Green and Myerson 2004; Cardinal 2006; Phillips, Walton et al. 2007; Rangel, Camerer et al. 2008). To do this, organisms formulate representations of potential courses of action as well as the internal and external states at the time of action. The costs and benefits of behavior based on these representations are evaluated and an action is selected. Once the action is performed the organisms compares the value of the outcome with the previous representations to update learning about the given situation (Rangel, Camerer et al. 2008). This type of value-based decision making can be modeled in humans and animals by exposing organisms to situations in which there is a choice between rewards of different value. For example, humans (Coffey, Gudleski et al. 2003; Green and Myerson 2004; Hariri, Brown et al. 2006; Prévost, Pessiglione et al. 2010) and animals (Cardinal, Pennicott et al. 2001; Roesch, Calu et al. 2007; Roesch, Singh et al. 2009; Day, Jones et al. 2010) show similar patterns of choice behavior based on the time the organism spends waiting for the reward. In these choice tasks, subjects are presented with smaller rewards presented sooner and larger rewards presented later and reward value is discounted by the time to reinforcement. Animals and humans show similar patterns of behavior, choosing the larger option less often as the delay to reward increases. Similar patterns of discounting are also seen when organisms must choose between rewards based on the cost of behavior (i.e. the effort required to obtain reinforcement (Floresco and Whelan 2009; Day 2010; Day, Jones et al. 2010; Gan, Walton et al. 2010; Prévost, Pessiglione et al. 2010)).

Risky decision making is a third type of value-based decision making which is impaired in several psychiatric disorders including gambling and drug addiction (Crews and Boettiger 2009). Risky decisions have been modeled in humans and animals using gambling paradigms in which organisms are given choices between “playing it safe” for a smaller reward delivered 100% of the time or “taking a risk” for a larger reward delivered with less certainty. Within this framework, organisms evaluate expected reward value that can be modeled with simple economic utility theory. One such metric, termed “reward value”, can be calculated by multiplying the explicit reward magnitude by the probability of reward delivery (Rangel, Camerer et al. 2008). For example, if a rat is expecting to get 2 sucrose pellets 50% of the time for performing a specific behavior, the expected value or net utility of the behavior is 1 sucrose pellet ($2 \times 0.5 = 1$). Following this logic, organisms can optimize choices by picking responses with the highest expected value. However, in some cases animals will be presented with situations in which they can choose between one or more options and the expected value of either performance is equal (e.g., rats are given a choice between 1 sucrose pellet 100% of the time versus 2 sucrose pellets 50% of the time). In these situations, the internal preferences about risk-taking and reward probability and magnitude determine the subjective value that organisms associate with the different options (Kable and Glimcher 2007). This is opposed to the objective or explicit reward value that is associated with the characteristics of the reinforcer itself (e.g., 4 pellets is always more valuable than 1 pellet). Animals display similar discounting to humans in that larger rewards are chosen less often as the probability of receiving the reward decreases (Green and Myerson 2004; Cardinal and Howes 2005; St Onge and Floresco 2008; Floresco and Whelan 2009; St. Onge, Chiu et al. 2010; St. Onge and Floresco 2010). Importantly, all value-based decision making

models incorporate situations in which organisms have choices between options that vary based on the different aspects of reward value. Both the explicit characteristics and subjective representations of reward value influence choices towards the most valuable options available.

Value-based decision making and the mesolimbic dopamine system

Recent research has employed several models of complex value-based decision making and has implicated the NAc and its afferent dopaminergic projections to the NAc in this behavior. Mesolimbic DA neurons encode information about reward value based on several factors including reward probability, delay, magnitude, and effort requirements to inform the organism of the value of options available (Fiorillo, Tobler et al. 2003; Tobler, Fiorillo et al. 2005; Roesch, Calu et al. 2007; Day, Jones et al. 2010). Therefore, perturbations in this system are expected to result in aberrant decision making. As such, lesions of the NAc induce risk-averse behaviors in a probabilistic reinforcement task (Cardinal and Howes 2005) and impulsive choices in a delay discounting task (Cardinal, Pennicott et al. 2001), even when these choices were less advantageous. Further, systemic administration of dopaminergic drugs alters risk-taking behavior and effortful decision making (St Onge and Floresco 2008; Floresco and Whelan 2009; St. Onge, Chiu et al. 2010) suggesting a role for the DA system in reward evaluation. Taken together, these results suggest that DA release in the terminal region of the NAc is critical in the mediation of appropriate decision making, and disruption of this circuitry is expected to alter decision making. For example, pharmacological disruption and 6-OHDA lesions of DA neurons in the NAc impairs the evaluation of costs associated with performing behaviors (Salamone, Wisniecki et al. 2001; Salamone, Salamone et al. 2002; Salamone, Correa et al. 2005;

Adriani, Boyer et al. 2009) while overexpression of the DA transporter in the NAc increases risk-taking behavior and impulsive choices. Recent evidence has also suggested that the mesolimbic DA projections (Roesch, Calu et al. 2007) and terminal DA release in the NAc (Day, Jones et al. 2010) encode the value of the options available during decision making tasks which functions to bias decisions towards the higher value options.

While previous value-based decision making research has implicated the NAc and mesolimbic DA function in this behavior, phasic DA signaling in the NAc during risky decision making based on subjective value remains unclear. Here, we investigated the role of phasic DA signaling during risky decision making in which animals were given the option to choose between a smaller certain and larger uncertain reinforcer. Fast-scan cyclic voltammetry was employed to evaluate how phasic DA signaling in the NAc encodes information about reward value based on risk. Importantly, the expected value of each behavioral response in the task was equal such that across the training session animals received the same amount of reward for taking risks or playing it safe. Therefore, we were able to examine the role of phasic DA signaling in decision making based on subjective value.

CHAPTER 2

METHODS

Animals

Male Sprague Dawley rats ($n = 8$ animals with 11 recording locations for NAc core and $n = 3$ animals with 3 recording locations for NAc shell, Harlan Sprague Dawley, Indianapolis, IN) aged 90-120d and weighing 275-350g were used as subjects and individually housed with a 12:12 light:dark cycle. All experiments were conducted between 9:00 am and 5:00 pm. Bodyweights were maintained at no less than 85% of pre-experimental levels by food restriction (10-15g of Purina laboratory chow each day, in addition to approximately 1g of sucrose consumed during daily sessions). This regimen was in place for the duration of behavioral testing, except during the post-operative recovery period when food was given *ad libitum*. All procedures were approved by the University of North Carolina at Chapel Hill Institutional Animal Care and Use Committee.

Behavioral Training

Behavioral training and recording sessions were conducted in 43 X 43 X 53cm Plexiglas chambers housed in a sound-attenuated cubicle (Med Associates, St Albans, VT). One side of the chamber had 2 retractable levers (Coulbourn Instruments, Allentown, PA) located 17cm apart. A stimulus light was located 6cm above each lever. A speaker to deliver white noise (80db) was located 12cm above the floor on the opposite wall. A houselight

(100mA) was mounted 6cm above the speaker. Forty-five mg sucrose pellets were delivered to the food receptacle located equidistantly between the retractable levers. During training sessions 1-5, both levers were extended and both cue lights above each lever were illuminated. Responding on either lever was reinforced on a continuous schedule of reinforcement (fixed ratio 1, FR1) and resulted in the delivery of a 45mg sucrose pellet. A maximum of 100 reinforcers (50 reinforcers per lever) were available. Next, training on the risky decision making task began (see figure 1a for schematic of task and figure 1b for session timeline). During the first 10 sessions rats were trained in a task that involved 3 types of contingencies (30 trials each) intermixed within 90 total trials per session. At this stage of training, only safe options were available to train rats to press each lever (i.e., risk trials were not included). The first two trial types were classified as forced-choice trials. For one trial type, a single cue light was illuminated for 5s over one lever followed by extension of both levers. Responses on the lever under the illuminated cue light within 15s were reinforced with 1 sucrose pellet. During the other forced-choice trial type, the other cue light over the second lever was illuminated (5s), followed by extension of both levers and responses (15s) on the corresponding lever were reinforced as above. Responses on the unsignaled lever were counted as “errors” and resulted in the houselight being extinguished for the remainder of the trial period and the absence of reward delivery. During the third trial type (termed free-choice trials), both cue lights were illuminated (5s), both levers were extended, and responses on either lever (15s) were reinforced with a single food pellet. Following each lever press during any trial type, the levers were retracted and rewards were immediately delivered. Thus, during the first 10 training sessions the reward contingencies for each option were identical (FR1 schedule of reinforcement for one 45mg sucrose pellet). This was done to

allow animals to fully learn the predictive associations of the cue lights before the reward contingencies were altered. Furthermore, this ensured that there would be no bias in response allocation as a result of differential learning between the two levers.

Following acquisition of sessions 1-10, the reward contingency on one of the levers (counterbalanced across animals) was altered to reflect the risky decision making task. As shown in Figure 1a, on forced-choice safe trials (top) a cue light was illuminated for 5 s followed by extension of both levers into the chamber. A response on the lever situated under that cue light resulted in one 45mg sucrose pellet delivered 100% of the time to a centrally located foodcup. Responses on the other lever were counted as errors, not rewarded, terminated the houselight, and ended the trial. On forced-choice risk trials (figure 1a, middle) the other cue light was presented for 5 s before extension of the lever beneath it. A response on that lever resulted in 2 sucrose pellets delivered 50% of the time to the foodcup. Responses on the other lever were counted as errors as above. On free-choice trials (figure 1a, bottom) both cue lights were illuminated and rats could select either option. As such, presses on the “safe lever” during forced- and free-choice trials resulted in the delivery of one 45 mg sucrose pellet 100% of the time. Presses on the “risk lever” during forced- and free-choice trials resulted in the delivery of two 45 mg sucrose pellets 50% of the time. Several other studies have used similar reward contingencies in which animals choose between smaller certain and larger uncertain rewards to model risk-taking behavior (Cardinal and Howes 2005; St Onge and Floresco 2008; Floresco and Whelan 2009; Roitman and Roitman 2010). Using this reinforcement contingency the expected value of each lever was equal such that presses on one lever were not more advantageous than presses on the other lever across the entire session. Animals were trained on the risky decision making task for an

additional 10 training sessions prior to the completion of electrochemical measurements (see below).

Surgery

Following behavioral training rats were surgically prepared for voltammetric recordings as previously described (Phillips, Robinson et al. 2003). Animals were anesthetized with a ketamine hydrochloride (100mg/kg) and xylazine hydrochloride (10mg/kg) mixture (intramuscular) and placed in a stereotaxic frame. A guide cannula (Bioanalytical Systems, West Lafayette, IN) was positioned dorsally to the NAc core (1.3 mm anterior, 1.3mm lateral from bregma). An Ag/AgCl reference electrode was placed contralateral to the stimulating electrode in the left forebrain. The bipolar stimulating electrode (Plastics 1 Inc., Roanoke, VA) was placed dorsally to the VTA (5.2 mm posterior, 1.0 mm lateral from bregma and 7 mm ventral from dural surface). Stainless steel skull screws and dental cement were used to secure all items. A detachable micromanipulator containing a glass-sealed carbon-fiber electrode (75-100 μ m exposed tip length, 7 μ m diameter, T-650; Amoco, Greenville, SC) was inserted into the guide cannula and the electrode lowered into the NAc. The bipolar stimulating electrode was lowered in 0.2 mm increments until electrically evoked DA release was detected at the carbon-fiber electrode in response to an electrical stimulation train (60 biphasic pulses, 60Hz, 120 μ A, 2 ms per phase). The stimulating electrode was then fixed with dental cement and the carbon-fiber electrode was removed.

Fast-scan cyclic voltammetry

Following surgery, animals were allowed to recover to presurgery body weight for one week. Food restriction was resumed to increase motivation during behavioral performance. Animals were retrained on the behavioral task for 4 sessions or until they reached presurgery levels of performance. DA concentration changes during behavior were assessed using fast-scan cyclic voltammetry as previously described (Day, Roitman et al. 2007). A new carbon-fiber electrode, housed in the micromanipulator, was lowered into the NAc core or shell and was used to measure DA changes during task performance. The potential of the carbon-fiber electrode was held at -0.4V versus the Ag/AgCl reference electrode. Voltammetric recordings were made every 100 ms by applying a triangular waveform that drove the potential to +1.3V and back to -0.4V at a scan rate of 400V/s. Application of the triangular waveform results in the oxidation and reduction of chemical species that are electroactive within this potential range (including DA), producing a measurable change in current at the carbon-fiber. Following equilibration in the brain (typically 20-30 minutes to reduce current drift at the electrode), DA release was electrically evoked by stimulating the VTA using a range of stimulation parameters (2-24 biphasic pulses, 20-60 Hz, 120 μ A, 2 ms per phase) to make sure that the carbon fiber electrode was placed close to DA release sites and to create a training set for principal component analysis. Animals then underwent task performance and electrochemical recordings were made continuously with 100 ms temporal resolution. Following termination of the behavioral session, VTA stimulation was repeated to verify the stability of the electrode and ensure that the location of the electrode still supported DA release.

Signal identification and separation

Stimulation of the VTA leads to two well-characterized electrochemical events: an immediate but transient increase in DA and a delayed but longer lasting basic pH shift. To separate these signals a training set was constructed from representative, background-subtracted cyclic voltammograms for DA and pH, as previously described (Heien, Khan et al. 2005). The background period (500msec) was obtained at the minima for the DA signal 5 sec before event onset. This training set was used to perform principal component regression on data collected during the behavioral session. Principal components were selected such that at least 99% of the variance in the training set was accounted for by the model. All data presented here fit the resulting model at the 95% confidence level. After use, carbon-fiber electrodes were calibrated in a solution of known DA to convert observed changes in current to differential concentrations.

Data analysis

All behavioral events (cue onset and offset, lever extension and retraction, lever presses and reward delivery) and electrochemical measurements were recorded. The number of errors and reward pellets received were evaluated using a one-way repeated measures ANOVA with Tukey's HSD post-hoc test. Response allocation on choice trials were calculated as 3 day averages of lever press behavior on free-choice trials (final two training sessions and the recording session or the final three training sessions before switching the reward contingencies). Animals were required to have at least 60% responses on the preferred lever during the recording session to be included in the data analysis. Using this criteria, 7 animals with 10 recording locations were included in the behavioral and electrochemical data analysis. Behavioral preferences were compared using an unpaired two-

tailed *t*-test to compare responses to chance level and paired two-tailed *t*-tests to compare preferred versus nonpreferred responses. Response latencies during the recording session were calculated by taking the average latency to lever press during forced-choice trials on the preferred and nonpreferred lever. Trials in which the response latency was greater than 10 seconds were excluded from analysis. The average response latencies were compared using a paired two-tailed *t*-test.

Phasic changes in extracellular DA concentration during the task were assessed by aligning DA concentration traces to relevant behavioral events (specifically cue presentation and lever extension). Group increases or decreases in NAc DA concentration from baseline in response to cue presentation were evaluated separately for each trial type (preferred, nonpreferred, and choice) using a one-way repeated measures ANOVA with Dunnett's correction for multiple comparisons. This analysis compared the baseline mean DA concentration to each data point (100msec bin) obtained within 2s following the cue presentation. The baseline DA concentration was the average DA concentration for 5s prior to cue onset for each animal.

In order to determine the relationship between cue evoked DA release and lever press behavior a correlation analysis was performed. The correlation between DA release in the NAc core and lever press behavior was performed by taking the ratio of the peak DA concentration within 2s of cue presentation for the risk and safe cues compared to the ratio of risk versus safe lever presses on free-choice trials. The ratio of DA signaling was calculated by taking the peak DA for the risk cue and dividing it by the total peak DA for the risk and safe cue presentation. The ratio for lever pressing was calculated by taking the total number of presses on the risk lever during free-choice trials and dividing by the total number of

presses during free-choice trials (sum of presses on the risk and safe lever). All animals were used for this analysis including the animal that did not display a behavioral preference. To assess the differential effects of the three cue types (preferred, nonpreferred, and choice) on DA release, the peak DA concentrations within 2 s following cue presentation were analyzed. The effects of the three cue types were evaluated using repeated measures ANOVA with Tukey's post-hoc tests. Comparisons were performed separately for data collected in the NAc core and shell. For comparison of DA signal on free-choice trials when animals pressed the preferred versus nonpreferred lever, one animal that never selected the nonpreferred option was removed to allow for proper statistical comparison and to ensure that results were not biased. Peak DA release during choice trials when the animal chose the preferred versus nonpreferred option were evaluated using a paired *t*-test. In order to confirm that the effect was not a result of an uneven number of trials, peak DA release during free-choice trials was also evaluated using a paired *t*-test in which trials that the rat pressed the preferred lever were randomly chosen to allow for equal number of preferred and nonpreferred trials. DA concentration during reward delivery was evaluated by examining the peak DA concentration within 2 seconds following lever extension for rewarded risk and safe trials and the lowest point within 2 seconds following lever extension for unrewarded risk trials. Again, data were evaluated using a repeated measures ANOVA with Tukey's post-hoc tests. All analysis were considered significant at $\alpha=0.05$. Statistical and graphical analysis were performed using GraphPad Prism 4.0 for Windows (GraphPad Software, La Jolla, CA), STATISTICA for Windows (StatSoft Inc. Tulsa OK) and Neuroexplorer for Windows version 4.034 (Plexon Inc., Dallas, TX).

Histology

Upon completion of the experiment, rats were deeply anesthetized with a ketamine/xylazine mixture (100 mg/kg and 10 mg/kg, respectively). To mark the placement of the electrode tip, a 150 to 250 μ A current was passed through a stainless steel electrode for 5 seconds using established procedures (Roitman, Stuber et al. 2004; Day, Roitman et al. 2007). Animals were decapitated and brains were removed and postfixed in 10% formalin. After postfixing and freezing, 40 μ m coronal brain sections were mounted on microscope slides. The specific position of electrodes was assessed by visual examination of successive coronal sections in comparison to visual landmarks and anatomical organization of the NAc represented in a stereotaxic atlas (Paxinos 2005). For the NAc core, multiple recordings in 3 animals were completed during different behavioral sessions in which the electrode was lowered to a new location.

CHAPTER 3

RESULTS

Behavioral performance on risky decision making task

Animals rapidly acquired the lever press response, performing maximal responding and consuming all rewards by the fifth pretraining session. Following pretraining, rats were able to learn the risky decision making task and discriminate between the cue types as evidenced by a significant reduction in the percentage of errors on forced safe or risk trials compared to session 1 ($F_{(24,144)} = 6.244$, $P < 0.0001$; figure 2a). On the final training session before voltammetry surgery animals averaged 0.5714 ± 0.7868 errors during the session. Furthermore, over the final 7 days of behavioral training, the number of errors was not significantly greater than 0. Concurrent with the reduction in errors, animals displayed a significant increase in the number of sucrose pellets received across all sessions compared to session 1 ($F_{(24,144)} = 3.092$, $P < 0.0001$; figure 2b). Taken together, these data show that behavioral performance improved with training such that by the recording session animals were performing the task optimally.

Response allocation on free-choice trials was evaluated to determine the behavioral preference of animals as they performed the risky decision making task. On free-choice trials animals displayed a significant preference for one contingency (safe or risky). Collapsed across contingency type (safe or risky), animals exhibited significantly greater presses on the preferred lever compared to the nonpreferred lever ($t(9) = 6.426$, $P = 0.001$) and pressed the

preferred lever significantly more than chance ($t(9) = 6.426$, $P = 0.001$; figure 2c). (One animal recorded in the NAc core did not display a behavioral preference during the recording session and thus is only included in the correlation analysis). Animals also showed a nonsignificant trend towards faster latency to press the preferred lever on forced choice trials during the recording session ($t(9) = 1.833$, $P = 0.1$; figure 2d). The preference data suggest that individual differences in risk-prone and risk-averse behaviors developed as the animals learned the task since lever preference was not observed early in training ($t(6) = 1.089$, $P > 0.05$, data not shown). Similar behavioral data were obtained for animals with recordings in the NAc shell. Specifically, those animals exhibited significantly more presses on the preferred versus nonpreferred lever ($t(2) = 8.362$, $P = 0.0140$) and pressed the preferred lever significantly faster than the nonpreferred lever ($t(2) = 5.094$, $P = 0.0364$ (figure 8a, b).

Reward-associated discriminative stimuli evoke phasic dopamine signals in the nucleus accumbens

The design of this task allowed us to investigate changes in NAc DA during the presentation of reward predictive cues and the delivery of expected and unexpected rewards. Consistent with previous reports reward-predictive cues in all trial types evoked the strongest increase in phasic DA release in both the NAc core (Stuber, Klanker et al. 2008; Day, Jones et al. 2010; Gan, Walton et al. 2010) and shell (Day, Jones et al. 2010) (see figure 7 for histological verification of electrode placements in the NAc core and shell). Representative single trial changes in DA concentration occurring during task performance are shown for two animals in figure 3a and figure 4a. The strongest increases in DA were typically observed following cue presentation during all three trial types for both risk-averse (figure

3a) and risk-prone (figure 4a) animals. A second lower concentration increase in DA was also observed during the reward delivery period. Figure 3b and 4b shows individual trial data as well as average changes in DA concentration for a risk-averse (figure 3b) and risk-prone (figure 4b) animal during the entire behavioral session. This demonstrates that throughout the training session cue presentations and reward deliveries evoked phasic increases in NAc DA concentration. In the next section, I examine how changes in DA reflect the preferred option of the animal (risk prone versus risk averse).

Cue-evoked dopamine signals within the nucleus accumbens core reflect reward preferences

Cue-selectivity in DA release was determined by comparing the peak DA release between the 3 trial types for each individual animal. On forced-choice trials, cue-evoked DA release in the NAc core tracked the subjective value of future rewards as DA signaling scaled with each animal's individual preference. For example, figure 3 shows DA release dynamics during the task for a single risk-averse animal. Cue-evoked DA release was higher for cues predicting safe (center) compared to risk (left) trials. This animal preferred the safe option and DA release during free-choice trials (right) was similar to that observed during forced-choice safe trials (center). Figure 4 shows a complementary pattern of DA concentration changes in a risk-prone animal, with larger increases in DA occurring for cues that predicted risky rewards. This pattern was conserved across all animals with recordings in the NAc core, as there was a significant correlation between the animal's risk taking behavior and the differential DA release ($r^2 = 0.7118$, $P = 0.0011$; figure 5).

Overall changes in DA concentration recorded at all locations in the NAc core for animals displaying a behavioral preference (n=7 animals with 10 recording locations) are

shown in figure 6a, aligned to cue onset. Specifically, cue onset occurred at time 0s and cue offset/lever presentation occurred at time 5s. Repeated measures ANOVA revealed that cue presentation in each trial type significantly increased DA concentration (Preferred: $F_{(20,9)} = 16.72$ $P < 0.0001$, Nonpreferred: $F_{(20,9)} = 28.39$ $P < 0.0001$, Choice: $F_{(20,9)} = 24.69$ $P < 0.0001$). Further, a one-way repeated measures ANOVA comparing the peak DA evoked on each trial type (maximum DA concentration within 2s of cue onset) showed that the amount of cue-evoked DA release varied depending on the type of cue presented ($F_{(2,9)} = 9.479$, $P = 0.0015$; figure 6b). Post hoc analysis revealed a significant difference in DA concentration for forced-choice trials during the presentation of the preferred versus nonpreferred cue (Tukey's *post hoc* test, $P < 0.01$; figure 6b). Further, during free-choice trials cue evoked DA release scaled to the preferred option. Specifically, peak DA concentration was similar for cues predicting the free-choice and preferred ($P > 0.05$; figure 6b) but not the nonpreferred option ($P < 0.01$, figure 6b).

Cue-evoked DA could be functioning to signal the most valuable option available or the value of the option that is eventually chosen. In order to distinguish between these, DA release events were also quantified during free-choice trials when the animal chose the preferred versus nonpreferred option. This analysis revealed that DA signaling on free-choice trials encoded the most valuable option available rather than what the animal subsequently chose. Specifically, there was no significant differences in DA release in the NAc core on free-choice trials when animals chose the preferred versus nonpreferred option ($t(8) = 0.8858$, $P = 0.4015$ for all choice trials and $t(8) = 2.089$, $P = 0.0701$ when number of trials were equal). DA signals on free-choice nonpreferred trials were also not significantly different from cue-evoked DA release on forced-choice preferred trials ($t(8) = 0.602$, $P =$

0.5624), suggesting that the DA signal encodes information about the most valuable options available during decision making even when the nonpreferred option was chosen. Finally, value encoding of DA signaling was subregion specific as no significant differences were observed in peak DA concentrations across cue types in the NAc shell despite a strong behavioral preference ($F_{(2,2)} = 4.931$, $P = 0.0833$; figure 8c and d).

Dopamine signaling in the nucleus accumbens core tracks prediction errors during risky decision making

It has been postulated by Schultz and colleagues that midbrain dopamine neurons encode information about stimulus-reward associations thus functioning as a learning signal (Schultz, Dayan et al. 1997). In these models, dopamine cell firing (examined by electrophysiological recording of dopamine neurons) is believed to provide a ‘prediction error’ signal that compares expected outcomes with actual outcomes (Schultz, Dayan et al. 1997). Unexpected rewards produce brief synchronous bursts among dopamine neurons (termed positive prediction error) while fully predicted rewards typically evoke little or no phasic activity. Further, if an expected reward is omitted, dopamine neurons exhibit a pause in cell firing, termed a negative prediction error (Schultz, Dayan et al. 1997).

The design of our behavioral task allows us to examine if DA release in the NAc core encodes prediction errors during risky decision making. To do this we evaluated DA release during a 2s period following lever presentation during forced-choice safe versus risk trials. The risk trials were separated into rewarded and unrewarded risk trials. Safe trials function as a situation in which rewards are perfectly predicted. Risk trials function as a situation where reward presentation and omission was unexpected. Figure 6c shows the timecourse of DA

release following lever extension in the NAc core. Unexpected reward delivery or omission during risk trials resulted in increases or decreases, respectively, in phasic DA concentration in the NAc core compared to perfectly predicted safe trials. Specifically, we observed a main effect of reward delivery on DA concentration ($F_{(2,9)} = 30.31$, $P < 0.0001$; figure 6d). Post hoc tests revealed that during forced-choice risk trials, reward presentation resulted in a significant increase in DA concentration compared to forced-choice safe trials (Tukey's *post hoc* test, $P < 0.001$; fig 6d). Thus, when reward was presented, an increase in DA release was observed consistent with a positive prediction error (Schultz, Dayan et al. 1997; Schultz 1998). Conversely, the omission of reward during forced-choice risk trials resulted in significantly less DA release compared to forced-choice safe trials ($P < 0.05$; figure 6d), consistent with the signaling of a negative prediction error (Schultz, Dayan et al. 1997; Schultz 1998).

CHAPTER 4

DISCUSSION

In the present study, DA release was measured in two subregions of the NAc while animals performed a risky decision making task. Importantly, the design of the task allowed us to assess how reward predictive cues associated with risky decisions affected patterns of DA release. We show that when given a choice animals develop individual preferences. They choose to either “play it safe” for a certain smaller reward, or to “take risks” for a larger, but uncertain reward. Reward predictive cues elicited an increase in DA concentrations in the NAc core that tracked individual preferences in the risky decision making task. That is, cues that predicted the preferred reward contingency evoked higher DA release than cues that predicted the nonpreferred option.

Further, the design of the experiment also allowed us to evaluate how DA release in the NAc core encodes information about the unexpected presentations and omissions of reward. As proposed by Schultz and colleagues (Schultz, Dayan et al. 1997), violations in the expectation of reward result in errors in prediction which function as a teaching signal for the animal. This teaching signal thereby increases or decreases the prediction of future outcomes and is thus believed to be a neural substrate for learning (Schultz, Dayan et al. 1997). Schultz and colleagues developed this model by examining the activity of mesolimbic dopamine neurons during electrophysiological recording techniques in awake, behaving monkeys. They showed that DA neurons in the VTA and substantia nigra encode a reward prediction error signal in which cues that predict rewards evoke phasic increases in firing rate while fully

expected rewards do not alter DA activity. Further, violations in these reward predictions alter DA activity, as unexpected reward delivery increases DA cell firing while unexpected reward omission decreases DA activity (Schultz, Dayan et al. 1997). Results from the current study support the theory that DA neurons that project to the NAc encode prediction errors. Specifically, increases in DA concentration were observed in the NAc core for unexpected reward deliveries and decreases in DA concentration occurred during unexpected reward omissions (Figure 6). Since our laboratory has previously shown that rapid dopamine release in the NAc is a direct result of changes in burst firing of DA neurons (Sombers et al., 2009), our findings are consistent with the reward prediction model by Schultz and colleagues.

Electrophysiological studies of DA cell firing show that cue-evoked DA signaling is sensitive to a number of features of upcoming reward such that cues which predict larger, immediate, or more probable rewards elicit larger spikes in DA neural activity than cues that predict smaller, delayed or less probable rewards (Fiorillo, Tobler et al. 2003; Tobler, Fiorillo et al. 2005; Roesch, Calu et al. 2007). This type of processing by DA neurons is hypothesized to be critical for decision making as it functions to broadcast information about reward value to striatal circuits that enable animals to maximize behaviors (Roesch, Calu et al. 2007; Day, Jones et al. 2010). Reward value is discounted by several “costs” including physical effort, time delays, probability, or aversive consequences, which animals must overcome to optimize choices (Phillips, Walton et al. 2007).

Using this framework, value-based decision making can be explained through basic utility functions as a “cost-benefit analysis.” Thus animals must evaluate the behavioral costs that discount reward value to determine if a behavior is beneficial. In order to do this, animals have an intrinsic “threshold” for behavior such that options that fall below the “cost

threshold” are deemed worthwhile, while options that fall above the cost threshold are deemed not worthwhile and will be rejected. It has been hypothesized that value encoding by the DA system is used to monotonically set the “cost threshold” (or breakpoint) beyond which the net outcome is no longer worthwhile. With this model, phasic DA activity in the NAc core signals information about the value of future rewards, and if this prediction is below the cost threshold the behavior is deemed worthwhile and the animal will perform the task. Conversely, options that are predicted to have low value will evoke much less DA release, and thus this signal of future value will fall above the cost threshold and will be deemed not worthwhile (Phillips, Walton et al. 2007).

Based on this model, animals with a hyperactive DA system will be able to overcome higher costs of behavior to perform a given task better than control animals. Conversely a depleted DA system results in animals that are not able to overcome the cost expenditure to obtain rewards compared to controls. Empirical evidence has supported this model and has shown that DA depletion in the NAc core selectively retarded responding when the effort to obtain rewards increased (i.e. increased lever pressing). When response costs were low, control and DA depleted animals performed similarly as the value of the response was well below the cost threshold in both groups of animals. However, as response costs increased, dopamine depleted animals began to stop selecting the high cost reinforcer as the reward value was now above the depleted animals cost threshold while control animals were able to complete the task (Salamone, Wisniecki et al. 2001; Phillips, Walton et al. 2007).

The “cost-benefit analysis” hypothesis of DA function has previously been used to explain if an animal will perform a given behavior based on the expected value of performance. However, this model can also be applied to a situation in which animals are

given concurrent choices with different values. Rather than serving as a threshold for performing a behavior or not, the DA system can be functioning to compare how two separate behaviors relate to the cost threshold and thus inform animals of which option is more worthwhile to guide choice behaviors. In support, evidence has shown that DA neural activity and terminal release encode information about the best available option when animals are given a concurrent choice of options with different explicit value, irrespective of what the animal actually chooses (Roesch, Calu et al. 2007; Day, Jones et al. 2010). The present data fit with this framework, as subjective value was encoded by the phasic DA release in the NAc core. Specifically, DA release on free-choice trials encoded the animals preferred, or most valuable, reward contingency even when the animal chose the less valuable option. Thus, DA release in the NAc core may play a key role in the evaluation of different responses and may thereby function to promote action selection.

The mesolimbic DA system appears to provide information on the best available option to an organism, however animals do not always perform optimally. Therefore, other components of the corticolimbic reward circuitry, such as the PFC, BLA, or NAc core contribute to the promotion of appropriate behavior. For example, when subjects were presented with cues that predicted rewards of different value based on differing probability and magnitude, neural activity in these corticolimbic regions encoded different components of reward value. For example, activity in the ventral striatum encoded information about the expected value of future reward as there was more activity for higher probability and magnitude rewards (Tobler, O'Doherty et al. 2007). Further, when people were presented with rewards presented after varying delays, ventral striatum activity correlated with each individuals subjective value (Kable and Glimcher 2007). Similarly, cue-activated neurons in

the NAc encode the future action based on reward value, suggesting that the NAc may promote action selection based on subjective value (Samejima, Ueda et al. 2005; Roesch, Singh et al. 2009; Day, Jones et al. 2011). Conversely, PFC activity in humans (specifically the orbitofrontal cortex (OFC) subregion) encoded information about reward uncertainty rather than expected value and action selection when presented with cues predicting reinforcers of differing probabilities and magnitudes. Further, the encoding of reward uncertainty was correlated with each individual's risk attitudes (Tobler, O'Doherty et al. 2007). Similarly, OFC neural responses to reward delivery were correlated with individual differences in risk-taking behavior (Roitman and Roitman 2010) suggesting that the PFC encodes information about intrinsic attitudes that influence reward value to bias decisions. Finally, the BLA is a structure that is critical for maintaining the representation of future reward value during behavior which is essential for maintaining appropriate responding (Winstanley, Theobald et al. 2004).

As such, each structure of the corticolimbic reward circuitry has a specific role in encoding reward value which mediates decision making, and thus damage to the different structures results in divergent impairments in risky decision making behavior. For example, inactivation of a subregion of the PFC disrupts risky decision making behavior, leading animals to become risk prone or risk averse, depending on which risk contingency the animal was first exposed to. This suggests animals were unable to update information on the changes in reward probability in each trial block to make appropriate choices (St. Onge and Floresco 2010). Conversely, BLA inactivation affected decision making behavior by biasing animals towards safe options during risky decision making (Ghods-Sharifi, St. Onge et al. 2009) suggesting that animals could not maintain the representation of reward in its absence and

thus biasing animals away from risky choices. Finally, inactivation (Stopper and Floresco 2011) or lesion (Cardinal and Howes 2005) of the NAc biased animals towards the safe option, even when this was less advantageous. These findings suggest that the NAc is critical for encoding expected value which promotes action selection. Collectively, these findings suggest that each neural substrate of this circuit play different roles in mediating appropriate decision making, and more specifically, the transfer of information between these structures and to the NAc is critical for risky decision making. As such, signaling from the PFC or BLA may override the value signaling of the mesolimbic DA system to bias responses towards a less valuable option which may explain why rats sometimes choose the nonpreferred lever (as shown here) or the less valuable option (Day, Jones et al. 2010; Day, Jones et al. 2011).

Although all cues evoked DA release in both the core and shell of the NAc in the present study, DA transmission of value information was restricted to the NAc core. These results suggest that reward prediction is signaled in the NAc shell independently of reward valuation related to risky decision making behavior. Several lines of research have demonstrated that DA transmission in the NAc core and shell signal different aspects of future rewards, including the valence (appetitive versus aversive rewards), novelty, and value of rewards. Specifically, DA transmission in the NAc core has been shown to be an expression of motivational value and novelty of future rewards while DA transmission in the shell is an expression of reward valence (Corbit, Muir et al. 2001; Bassareo, De Luca et al. 2002). Recent studies have demonstrated hedonic valence encoding in the NAc shell showing increases in DA release for a sweet sucrose solution and decreases in DA for a bitter quinine solution (Roitman, Wheeler et al. 2008). Consistent with this idea, damage to the core versus shell of the NAc has dissociable effects on value-based decision making. For example,

inactivation of the core, but not the shell, impairs the ability to overcome high response costs (Ghods-Sharifi and Floresco 2010) resulting in animals choosing low effort options more often than controls, while lesions of the NAc core and not the shell lead to increased impulsivity in several impulsive choice tests (Pothuizen, Jongen-Rêlo et al. 2005). Finally, lesions of the NAc core result in risk averse behaviors (even when this is the less advantageous behavior (Cardinal and Howes 2005)) while inactivation of the NAc shell alone produce much less robust risk averse behavior than inactivation of the NAc core and shell together (Stopper and Floresco 2011). Disruption of NAc shell activity did not have substantial effects on any value based decision making tasks, suggesting that the NAc core is critical for encoding reward value that mediates decisions. As such, valence based DA signaling in the shell is not expected to be different when cues are presented that predict the same reward with different values as was observed here and during an effort based decision making task (Day, Jones et al. 2010).

The mesolimbic DA system is critical for adaptive reward seeking and decision making behavior, however it has also been implicated in maladaptive behaviors including drug addiction. Drug addiction involves the same circuitry as natural reward seeking and associative learning, as cues that are associated with drug taking also evoke phasic increases in DA in the NAc (Phillips, Stuber et al. 2003). A hallmark of drug addiction is increased risk-taking behavior and poor decision making (Jentsch and Taylor 1999; Robbins and Everitt 1999). As such, imbalances in the mesolimbic DA system may impair normal reward valuation processes and influence aberrant decision making behavior. For example, in rats that display trait impulsivity there is a significant reduction in dopamine D2/D3 receptor availability which is further correlated with subsequent increases in drug taking behavior

(Dalley, Fryer et al. 2007). Further, D2 receptor availability in the striatum is correlated with social dominance in male monkeys which is also related to cocaine intake. Specifically, subordinate monkeys show reduced D2 receptor availability compared to controls and also show increased drug taking behavior (Morgan, Grant et al. 2002). Alterations in the value encoding of mesolimbic DA system is a proposed mechanism of action which mediates aberrant decisions such as drug taking behavior (Schultz 2011). As discussed above, when cues become associated with natural rewards, NAc DA release in response to reward presentation decreases. However, addictive drugs produce changes in the mesolimbic DA neurons that result in increased excitability of these neurons to rewards and reward predictive cues (Jones and Bonci 2005). As such, it is hypothesized that drugs of abuse elicit increased DA release to both cues and rewards regardless of the prediction, thus always signaling a positive prediction error, or a reward that was better than predicted (Redish 2004; Schultz 2011). In turn, this would result in a situation in which drug reinforcers are always increasing in predictive value biasing decisions in favor of drug seeking behavior (Redish 2004). The mesolimbic DA system therefore is critical for encoding reward value that mediates appropriate value-based decision making, however imbalances in this encoding may result in maladaptive decisions related, for example, to drug addiction.

Figure 1 (a) Schematic representation of risky decision making task. See methods for detailed description of task. After training, NAc DA was measured with fast-scan cyclic voltammetry during a single 90-trial behavioral session. **(b)** Experimental timeline. Animals received 20 total training sessions before surgical implantation of the guide cannula above the NAc (each circle = 1 session). During the first 10 sessions of training the reward contingency for both the safe and risk option were identical (both levers rewarded with one 45 mg sucrose pellet 100% of the time) to ensure that animals learned the task without bias. During the next 10 sessions animals were trained as explained in **(a)**. Following surgery animals were reexposed to the task and underwent at least 4 days of retraining or until they returned to presurgical behavior and DA was recorded a single test session.

Figure 2 Behavior on risky decision making task. **(a)** Task performance revealed a significant reduction in the number of errors across training session compared to session 1. By the end of training the number of errors was not significantly different from 0. **(b)** A significant increase was observed in the number of pellets received compared to session 1. By the end of training animals earned the maximum number of pellets possible. **(c)** Response allocation on free-choice trials averaged across the final two training sessions and the recording session. Dashed line represents indifference point. Each animal showed a significant preference for one of the reward (safe or risk) options. **(d)** Response latency for forced-choice trials during the recording session. Animals showed a nonsignificant trend towards a decreased response latency for the preferred versus nonpreferred lever. All data are mean $\pm$ SEM. * $P < 0.05$; ** $P < 0.0001$; # $P = 0.1$.

Figure 3 Representative electrochemical data during individual trials for a risk-averse animal. **(a)** (Top) Three-dimensional representation (time x voltage x current) of electrochemical data during a single forced-choice risk, forced-choice safe, and free-choice trial and the corresponding DA concentration trace (bottom). Signal is aligned to cue onset (time 0s) and cue offset/lever extension (time 5s; colored bars along x axis). Cues evoked differential DA release that was correlated with animal's safe preference, evoking higher DA release for forced-choice safe and free-choice trials. **(b)** DA release averaged across all trials of the same type for the animal presented in **(a)**. Heat plots (top) represent individual trial data ordered with first trial on top. Signal is aligned to cue onset (time 0s) and cue offset/lever extension (time 5s; colored bars along x axis). DA trace (bottom) represents the average DA for all trials of each type. Across the entire training session, cues evoked differential DA release that was correlated with the animal's safe preference.

Figure 4 Representative electrochemical data during individual behavioral trials for a risk-prone animal. **(a)** Cue-evoked DA release for an individual forced-choice risk, forced-choice safe, and free-choice trial. Conventions follow figure 3a. Cues evoked differential DA release that was correlated with animal's risk preference, evoking higher DA release for forced-choice risk and free-choice trials. **(b)** Changes in DA across multiple trials for the animal presented in **(a)**. Conventions follow figure 3b. Across the entire training session, cues evoked differential DA release that was correlated with the animal's risk preference.

Figure 5 DA signaling correlates with lever press behavior for risk versus safe options. The x-axis is lever press behavior during free choice trials showing the risk preference of each animal. The ratio of lever pressing was determined by dividing the number of presses on the risk lever during free-choice trials by the total number of presses on free-choice trials (Risk/(Risk+Safe)). A ratio greater than 0.5 suggests that an animal is risk-prone while a ratio of less than 0.5 suggests an animal is risk-averse. A ratio equaling 0.5 suggests that an animal has no preference. The y-axis is the ratio of DA signaling for forced-choice risk and forced-choice safe trials. This ratio was calculated by taking the peak DA concentration within 2s of the cue presentation for risk and safe trials using the same formula as above (Risk/Risk+Safe). A ratio higher than 0.5 means that an animal had greater DA signaling during the risk cue compared to the safe cue while a ratio of less than 0.5 means that an animal had less DA signaling to the risk cue than the safe cue. A ratio of 0.5 means that DA signaling was equal across the two cue types.

Figure 6 DA release in the NAc core encodes (a, b) subjective value and (c, d) prediction errors during the risky decision-making task. (a) DA concentration on forced-choice trials (separated by each animal's behavioral preference) and free-choice trials aligned to cue onset (black bar, time 0s). (b) Peak cue-evoked DA concentration during a 2 s period following cue onset. Cue presentation on preferred and free-choice trials led to significantly larger increases in DA concentration than cue presentation on nonpreferred trials. (c) DA release during forced-choice risk trials separated by reward presentation or omission. Average lever press across animals at black triangle; line around black triangle shows minimum/maximum of average press. (d) DA concentration within 2 s following lever presentation. Unexpected presentation of reward during risk trials evoked a significant increase in DA compared to perfectly predicted safe trials while unexpected reward omission caused a significant decrease in DA concentration compared to safe trials. All data are mean $\pm$ SEM. * $p < 0.05$; ** $p < 0.01$.

Figure 7 Anatomical distribution of carbon fiber electrode placements. Coronal sections showing location of electrode tip during recording sessions in the NAc core and shell. Numbers to the right indicate distance anterior to bregma. Serial sections adapted from Paxinos and Watson (2005).

Figure 8 DA signaling in the NAc shell does not track subjective value during risky decision making. **(a)** Response allocation on free-choice trials averaged across the final two training sessions and the recording session for NAc shell animals. Dashed line represents indifference point. Each animal showed a significant preference for one of the reward (safe or risk) options. **(b)** Response latency for forced-choice trials during the recording session. Animals showed a significantly decreased response latency for the preferred versus nonpreferred lever. **(c)** DA concentration on forced-choice trials (separated by each animal's behavioral preference) and free-choice trials aligned to cue onset (black bar, time 0s). **(d)** Peak cue-evoked DA signal within 2 seconds of cue onset. There was no significant difference in cue-evoked DA concentration for the three cue types $P = 0.0833$. All data are mean $\pm$ SEM. $*P < .05$.

REFERENCES

- Adriani W, Boyer F, et al. (2009). Increased impulsive behavior and risk proneness following lentivirus-mediated dopamine transporter over-expression in rats' nucleus accumbens. *Neuroscience* 159(1): 47-58.
- Bassareo V, De Luca MA, et al. (2002). Differential Expression of Motivational Stimulus Properties by Dopamine in Nucleus Accumbens Shell versus Core and Prefrontal Cortex. *The Journal of Neuroscience* 22(11): 4709-4719.
- Berridge KC and Robinson TE (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews* 28(3): 309-369.
- Cardinal RN (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw* 19(8): 1277-1301.
- Cardinal RN and Howes NJ (2005). Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats. *BMC Neurosci* 6: 37.
- Cardinal RN, Pennicott DR, et al. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science* 292(5526): 2499-2501.
- Carelli RM (2000). Activation of accumbens cell firing by stimuli associated with cocaine delivery during self-administration. *Synapse* 35(3): 238-242.
- Carelli RM (2004). Nucleus accumbens cell firing and rapid dopamine signaling during goal-directed behaviors in rats. *Neuropharmacology* 47(Supplement 1): 180-189.
- Carelli RM and Deadwyler SA (1994). A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. *The Journal of Neuroscience* 14(12): 7735-7746.
- Carelli RM, Ijames SG, et al. (2000). Evidence that separate neural circuits in the nucleus accumbens encode cocaine versus "natural" (water and food) reward. *J Neurosci* 20(11): 4255-4266.

- Carelli RM, King VC, et al. (1993). Firing patterns of nucleus accumbens neurons during cocaine self-administration in rats. *Brain Research* 626(1-2): 14-22.
- Coffey SF, Gudleski GD, et al. (2003). Impulsivity and rapid discounting of delayed hypothetical rewards in cocaine-dependent individuals. *Exp Clin Psychopharmacol* 11(1): 18-25.
- Corbit LH, Muir JL, et al. (2001). The Role of the Nucleus Accumbens in Instrumental Conditioning: Evidence of a Functional Dissociation between Accumbens Core and Shell. *The Journal of Neuroscience* 21(9): 3251-3260.
- Crews FT and Boettiger CA (2009). Impulsivity, frontal lobes and risk for addiction. *Pharmacology Biochemistry and Behavior* 93(3): 237-247.
- Cromwell HC and Schultz W (2003). Effects of Expectations for Different Reward Magnitudes on Neuronal Activity in Primate Striatum. *Journal of Neurophysiology* 89(5): 2823-2838.
- Dalley JW, Fryer TD, et al. (2007). Nucleus accumbens D2/3 receptors predict trait impulsivity and cocaine reinforcement. *Science* 315(5816): 1267-1270.
- Day JJ, Jones JL, et al. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *European Journal of Neuroscience* 33(2): 308-321.
- Day JJ, Jones JL, et al. (2010). Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. *Biol Psychiatry* 68(3): 306-309.
- Day JJ, Jones, J.L., Carelli, R.M., (2010). Nucleus accumbens neurons encode predicted and ongoing reward costs. *European Journal of Neuroscience* In Press.
- Day JJ, Roitman MF, et al. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci* 10(8): 1020-1028.
- Day JJ, Wheeler RA, et al. (2006). Nucleus accumbens neurons encode Pavlovian approach behaviors: evidence from an autoshaping paradigm. *Eur J Neurosci* 23(5): 1341-1351.
- Fiorillo CD, Tobler PN, et al. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299(5614): 1898-1902.

- Floresco SB and Whelan JM (2009). Perturbations in different forms of cost/benefit decision making induced by repeated amphetamine exposure. *Psychopharmacology (Berl)* 205(2): 189-201.
- Gan JO, Walton ME, et al. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat Neurosci* 13(1): 25-27.
- Ghods-Sharifi S and Floresco SB (2010). Differential effects on effort discounting induced by inactivations of the nucleus accumbens core or shell. *Behavioral Neuroscience* 124(2): 179-191.
- Ghods-Sharifi S, St. Onge JR, et al. (2009). Fundamental Contribution by the Basolateral Amygdala to Different Forms of Decision Making. *J. Neurosci.* 29(16): 5251-5259.
- Goto Y and Grace AA (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat Neurosci* 8(6): 805-812.
- Green L and Myerson J (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychol Bull* 130(5): 769-792.
- Hariri AR, Brown SM, et al. (2006). Preference for Immediate over Delayed Rewards Is Associated with Magnitude of Ventral Striatal Activity. *The Journal of Neuroscience* 26(51): 13213-13217.
- Heien MLAV, Khan AS, et al. (2005). Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats. *Proceedings of the National Academy of Sciences of the United States of America* 102(29): 10023-10028.
- Hollander JA and Carelli RM (2007). Cocaine-associated stimuli increase cocaine seeking and activate accumbens core neurons after abstinence. *J Neurosci* 27(13): 3535-3539.
- Jentsch JD and Taylor JR (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology* 146(4): 373-390.
- Jones JL, Day JJ, et al. (2010). The basolateral amygdala differentially regulates conditioned neural responses within the nucleus accumbens core and shell. *Neuroscience* 169(3): 1186-1198.

- Jones S and Bonci A (2005). Synaptic plasticity and drug addiction. *Curr Opin Pharmacol* 5(1): 20-25.
- Kable JW and Glimcher PW (2007). The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10(12): 1625-1633.
- Kelley AE (2004). Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. *Neuroscience & Biobehavioral Reviews* 27(8): 765-776.
- Kim H, Sul JH, et al. (2009). Role of striatum in updating values of chosen actions. *J Neurosci* 29(47): 14701-14712.
- Mogenson GJ, Jones DL, et al. (1980). From motivation to action: Functional interface between the limbic system and the motor system. *Progress in Neurobiology* 14(2-3): 69-97.
- Morgan D, Grant KA, et al. (2002). Social dominance in monkeys: dopamine D2 receptors and cocaine self-administration. *Nat Neurosci* 5(2): 169-174.
- Nicola SM, Taha SA, et al. (2005). Nucleus accumbens dopamine release is necessary and sufficient to promote the behavioral response to reward-predictive cues. *Neuroscience* 135(4): 1025-1033.
- Nicola SM, Yun IA, et al. (2004). Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task. *J Neurophysiol* 91(4): 1840-1865.
- Nicola SM, Yun IA, et al. (2004). Firing of Nucleus Accumbens Neurons During the Consummatory Phase of a Discriminative Stimulus Task Depends on Previous Reward Predictive Cues. *Journal of Neurophysiology* 91(4): 1866-1882.
- Paxinos G (2005). The rat brain in stereotaxic coordinates. New York, Elsevier Academic Press.
- Phillips PE, Robinson DL, et al. (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.

- Phillips PE, Stuber GD, et al. (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* 422(6932): 614-618.
- Phillips PE, Walton ME, et al. (2007). Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology (Berl)* 191(3): 483-495.
- Pothuizen HHJ, Jongen-Rêlo AL, et al. (2005). Double dissociation of the effects of selective nucleus accumbens core and shell lesions on impulsive-choice behaviour and salience learning in rats. *European Journal of Neuroscience* 22(10): 2605-2616.
- Prévost C, Pessiglione M, et al. (2010). Separate Valuation Subsystems for Delay and Effort Decision Costs. *The Journal of Neuroscience* 30(42): 14080-14090.
- Rangel A, Camerer C, et al. (2008). A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9(7): 545-556.
- Redish AD (2004). Addiction as a computational process gone awry. *Science* 306(5703): 1944-1947.
- Robbins TW and Everitt BJ (1999). Drug addiction: bad habits add up. *Nature* 398(6728): 567-570.
- Roesch MR, Calu DJ, et al. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci* 10(12): 1615-1624.
- Roesch MR, Singh T, et al. (2009). Ventral Striatal Neurons Encode the Value of the Chosen Action in Rats Deciding between Differently Delayed or Sized Rewards. *Journal of Neuroscience* 29(42): 13365-13376.
- Roitman JD and Roitman MF (2010). Risk-preference differentiates orbitofrontal cortex responses to freely chosen reward outcomes. *European Journal of Neuroscience* 31(8): 1492-1500.
- Roitman MF, Stuber GD, et al. (2004). Dopamine operates as a subsecond modulator of food seeking. *J Neurosci* 24(6): 1265-1271.

- Roitman MF, Wheeler RA, et al. (2005). Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* 45(4): 587-597.
- Roitman MF, Wheeler RA, et al. (2008). Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nat Neurosci* 11(12): 1376-1377.
- Salamone, Salamone J, et al. (2002). Dopamine antagonists alter response allocation but do not suppress appetite for food in rats: contrast between the effects of SKF 83566, raclopride and fenfluramine on a concurrent choice task. *Psychopharmacology* 160(4): 371-380.
- Salamone JD, Correa M, et al. (2005). Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. *Curr Opin Pharmacol* 5(1): 34-41.
- Salamone JD, Wisniecki A, et al. (2001). Nucleus accumbens dopamine depletions make animals highly sensitive to high fixed ratio requirements but do not impair primary food reinforcement. *Neuroscience* 105(4): 863-870.
- Samejima K, Ueda Y, et al. (2005). Representation of action-specific reward values in the striatum. *Science* 310(5752): 1337-1340.
- Schultz W (1998). Predictive reward signal of dopamine neurons. *J Neurophysiol* 80(1): 1-27.
- Schultz W (2011). Potential vulnerabilities of neuronal reward, risk, and decision mechanisms to addictive drugs. *Neuron* 69(4): 603-617.
- Schultz W, Apicella P, et al. (1992). Neuronal activity in monkey ventral striatum related to the expectation of reward. *Journal of Neuroscience* 12(12): 4595-4610.
- Schultz W, Dayan P, et al. (1997). A neural substrate of prediction and reward. *Science* 275(5306): 1593-1599.
- Somers LA, Beyene M, et al. (2009). Synaptic overflow of dopamine in the nucleus accumbens arises from neuronal activity in the ventral tegmental area. *J Neurosci* 29(6): 1735-1742.

- St Onge JR and Floresco SB (2008). Dopaminergic Modulation of Risk-Based Decision Making. *Neuropsychopharmacology*.
- St Onge JR and Floresco SB (2008). Dopaminergic Modulation of Risk-Based Decision Making. *Neuropsychopharmacology* 34(3): 681-697.
- St. Onge J, Chiu Y, et al. (2010). Differential effects of dopaminergic manipulations on risky choice. *Psychopharmacology (Berl)* 211(2): 209-221.
- St. Onge JR and Floresco SB (2010). Prefrontal Cortical Contribution to Risk-Based Decision Making. *Cerebral Cortex* 20(8): 1816-1828.
- Stopper C and Floresco S (2011). Contributions of the nucleus accumbens and its subregions to different aspects of risk-based decision making. *Cognitive, Affective, & Behavioral Neuroscience* 11(1): 97-112.
- Stratford TR and Kelley AE (1999). Evidence of a Functional Relationship between the Nucleus Accumbens Shell and Lateral Hypothalamus Subserving the Control of Feeding Behavior. *Journal of Neuroscience* 19(24): 11040-11048.
- Stuber GD, Klanker M, et al. (2008). Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons. *Science* 321(5896): 1690-1692.
- Tobler PN, Fiorillo CD, et al. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307(5715): 1642-1645.
- Tobler PN, O'Doherty JP, et al. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J Neurophysiol* 97(2): 1621-1632.
- Waelti P, Dickinson A, et al. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412(6842): 43-48.
- Winstanley CA, Theobald DE, et al. (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J Neurosci* 24(20): 4718-4722.

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

Zellner MR, Kest K, et al. (2009). NMDA receptor antagonism in the ventral tegmental area impairs acquisition of reward-related learning. *Behav Brain Res* 197(2): 442-449.