

**THE INFLUENCE OF DRUGS OF ABUSE ON REWARD-BASED DECISION-
MAKING**

SCOTT WONG

Bachelor of Science, University of Lethbridge, 2014

A Thesis

Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

MASTER OF SCIENCE

Department of Neuroscience
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

THE INFLUENCE OF DRUGS OF ABUSE ON REWARD-BASED DECISION-
MAKING

SCOTT WONG

Date of Defence: August 25, 2016

Dr. A. Gruber Supervisor	Associate Professor	Ph.D.
-----------------------------	---------------------	-------

Dr. G. Metz Thesis Examination Committee Member	Professor	Ph.D.
--	-----------	-------

Dr. B. Kolb Thesis Examination Committee Member	Professor	Ph.D.
--	-----------	-------

Dr. D. Euston Chair, Thesis Examination Committee	Assistant Professor	Ph.D.
--	---------------------	-------

Dedication

To my family. Throughout my graduate journey you have always been there for me, congratulating me after my triumphs and encouraging me after my failures. You took an interest in my work, even when I couldn't explain it. You listened to my rants and always knew how to get me to look on the bright side. I'm grateful for all these things and without your love and support this would not have been possible.
Thank you.

And to Aaron, in the immortal words of Dr. Hubert Farnsworth, "I'm sciencing as fast as I can".

Abstract

Animals are thought to learn from reinforcement by a dopamine-dependent reward prediction error, based on the difference between the expected and actual value of a reward. Here we demonstrate that this learning signal can be manipulated by drugs of abuse (amphetamine and Δ -9-tetrahydrocannabinol), resulting in reduced loss sensitivity on a competitive binary choice task. This effect is distinct from global choice strategy, as the randomness of choice responding is unaffected by drug and animals with on-board Δ -9-tetrahydrocannabinol are still capable of flexibly tracking changing reward contingencies.

Furthermore, we demonstrate that lose-shift responding is increased following amphetamine sensitization. We propose this increase is indicative of drug-induced plastic changes facilitating a shift in behavioral control towards dorsolateral striatum. These results highlight the potential for drugs of abuse to alter reward-based decision-making and provide a novel measure of the shift in behavioral control towards dorsolateral striatum that is proposed to occur in addiction.

Acknowledgements

I would like to thank my supervisor, Dr. Aaron Gruber for his support and guidance during the last two years. His help has been invaluable not only in my work, but also in my progression towards becoming a better scientist. I would also like to thank the members of my supervisory committee, Dr. Bryan Kolb and Dr. Gerlinde Metz. Their experience and suggestions have greatly aided me during my graduate work. I also want to acknowledge the help I have received from various members of the Gruber lab. I am greatly appreciative of their assistance and could not have completed my Master's degree without it.

Finally, thank you to my friends and family, who have supported me in my graduate journey and put up with me as I prepared my thesis.

Table of Contents

Approval	ii
Dedication	iii
Abstract	iv
Acknowledgements	v
List of Tables	ix
List of Figures	x
List of Abbreviations	xi
1. General Introduction	1
1.1 Dopamine-dependent reward learning signals	1
1.2 Effects of drugs of abuse on dopamine transmission and prediction errors.....	2
1.3 Drug sensitization, synaptic plasticity, and decision-making	5
2. Opposing effects of acute and chronic d-amphetamine on decision-making in rats	7
2.1 Introduction	8
2.2 Methods	10
2.2.1 Subjects	10
2.2.2 Apparatus and choice task	10
2.2.3 Drug preparation and injections.....	11
2.2.4 Experiment 1: Acute effects of AMPH on the MP task	12
2.2.5 Experiment 2: Effects of AMPH sensitization on the MP task	12
2.2.6 Experiment 3: Effects of AMPH infusion into DLS and VS on the MP task ...	13
2.2.6.1 Surgery	13
2.2.6.2 Behavior testing	14
2.2.6.3 Histology	14
2.2.7 Behavior analysis.....	15
2.2.8 Tissue/diolistic bullet preparation and diolistic labeling.....	15
2.2.9 Confocal imaging and dendritic spine density analysis of DiI-impregnated tissue sections.....	16
2.3 Results	17
2.3.1 Experiment 1 (acute effects of AMPH on the MP task).....	17
2.3.2 Experiment 2 (Effects of AMPH sensitization on the MP task)	19
2.3.3 Dendritic spine density analysis	21
2.3.4 Experiment 3: Acute effects of AMPH intracerebral infusion on the MP task.	22
2.4 Discussion	24
2.5 Conclusion.....	30
3. Acute Δ -9-tetrahydrocannabinol reduces loss sensitivity in a binary choice task while preserving global behavioral flexibility	31
3.1 Introduction	32
3.2 Methods	34
3.2.1 Subjects	34
3.2.2 Behavior apparatus	35
3.2.3 Task 1: Matching Pennies (MP) Task.....	35
3.2.4 Task 2: Probability Reversal (PR) Task	36
3.2.5 Drug preparation and injections.....	37

3.2.6 Experiment 1: Acute effects of THC on the MP task.....	37
3.2.7 Experiment 2: Acute effects of THC on PR task	37
3.2.8 Behavior analysis.....	38
3.3 Results	39
3.3.1 Experiment 1 (Acute effects of THC on the MP task)	39
3.3.2 Experiment 2 (Acute effects of THC on the PR Task).....	41
3.4 Discussion	45
3.5 Conclusion.....	50
4. General Discussion	52
5. General Conclusion.....	56
6. References.....	57

List of Tables

Table 1. Group averages of number of trials per session and response entropy for AMPH experiments.24

Table 2. Comparison of behavioral measures influenced by THC in the Matching Pennies and Probability Reversal tasks..46

Table 3. Comparison of effects of systemic AMPH and THC on behavioral measures in the MP task.54

List of Figures

Figure 1. Matching Pennies schematic diagram and flow chart detailing a trial and loss-shift responding.....3
Figure 2: Coronal section of rat brain detailing striatal sub-regions.....4
Figure 3. Effects of acute AMPH administration on behavioral measures..... 19
Figure 4. Effects of AMPH sensitization on behavioral measures.20
Figure 5. Spine density quantification in DLS.....22
Figure 6. Histology and behavioral effects of AMPH infusion into DLS or VS.....24
Figure 7. Effects of acute THC administration on behavioral measures in the MP task...40
Figure 8. Representative data (240 trials over 4 blocks) of a rat’s feeder choices during the PR task.42
Figure 9. Session-averaged effects of acute THC in the PR task.43
Figure 10. Within-block effects of THC on behavioral measures in the PR task.....45

List of Abbreviations

AMPH	Amphetamine
ANOVA	Analysis of Variance
CB1	Cannabinoid Receptor 1
DAT	Dopamine Transporter
DiI	1,1'-Dioctadecyl-3,3,3',3'-Tetramethylindocarbocyanine Perchlorate
DLS	Dorsolateral Striatum
DMS	Dorsomedial Striatum
IFA	Impulsive Feeder Approach
IP	Intraperitoneal
LPFC	Lateral Prefrontal Cortex
MP	Matching Pennies
MSN	Medium Spiny Neuron
OFC	Orbitofrontal Cortex
PBS	Phosphate-Buffered Saline
PFA	Paraformaldehyde
PR	Probability Reversal
RPE	Reward Prediction Error
THC	Δ -9-Tetrahydrocannabinol
VS	Ventral Striatum
VTA	Ventral Tegmental Area
SN	Substantia Nigra

1. General Introduction

1.1 Dopamine-dependent reward learning signals

The ability of animals to alter their decisions in response to a changing environment confers a fitness advantage over less adaptable competitors. The neural mechanisms underlying such behavioral flexibility have been debated for more than a century. Classic work from Wolfram Schultz has provided evidence that the dopaminergic neurons of the midbrain may function in providing a reinforcement-based learning signal to achieve this (Schultz et al., 1997). Schultz observed that dopamine neurons in ventral tegmental area (VTA) fired when a monkey received an unexpected reward. After repeated pairings of this reward with a conditioned stimulus, dopamine neuron firing shifted to the presentation of the reward-predicting cue rather than the reward itself. Finally, when the conditioned stimulus was presented but no reward delivered, there was a pause in firing. These observations led him and others to propose that dopamine neurons provided a reward prediction error (RPE) signal. When greater than expected reward is received, there is an increase in dopamine neuron firing and a positive RPE. A reward that is equal in magnitude to what it expected elicits no changes in firing and is therefore a neutral RPE. The omission of reward or a smaller than expected reward reduces firing and elicits a negative RPE. Current reinforcement learning theory proposes that this mechanism provides feedback as to the outcome of choices to update internal representations of values associated with these choices. However, changes in phasic dopamine levels may also trigger other mechanisms that influence decision-making.

1.2 Effects of drugs of abuse on dopamine transmission and prediction errors

All known drugs of abuse (including cocaine, amphetamine, and Δ -9-tetrahydrocannabinol) increase dopamine levels in the brain by various mechanisms. For example, cocaine binds to the dopamine transporter (DAT) and prevents transmitter reuptake (Ritz et al., 1987). Similarly, amphetamine (AMPH) increases extracellular dopamine concentration by reversing DAT function (Fleckenstein et al., 2007). Conversely, Δ -9-tetrahydrocannabinol (THC) binds to cannabinoid receptors expressed on inhibitory afferents to midbrain dopaminergic neurons and by presynaptically reducing inhibition on these cells, it serves to increase dopamine release (Sañudo-Peña et al., 1999).

These effects on dopamine lead us to hypothesize that on-board drugs of abuse can manipulate RPE and therefore affect decision-making. Specifically we hypothesize the increase in dopamine at the synaptic cleft from drug action will attenuate the normal dip in dopamine that serves as a negative reward prediction error. Without this negative reinforcement signal, rats will be more likely to exhibit reduced loss sensitivity on reward-based decision-making tasks.

To assess reward-based choice behavior, the Gruber lab utilizes an uncued competitive binary choice task called Matching Pennies (MP). In this task rats initiate trials with a nosepoke into a central port, and then go to the left or right feeder to potentially receive a sucrose water reward. Loss sensitivity can be assayed by the prevalence of lose-shift responding, which is represented by rats on the current trial choosing not to return to the feeder that on the previous trial did not give a reward. Correspondingly, reduced loss sensitivity is represented by returning to a feeder which

did not provide a reward on the previous trial. (see Fig 1). Based on our hypothesis of drugs of abuse reducing loss sensitivity, we predict that rats with drug on-board will exhibit reduced lose-shift responding.

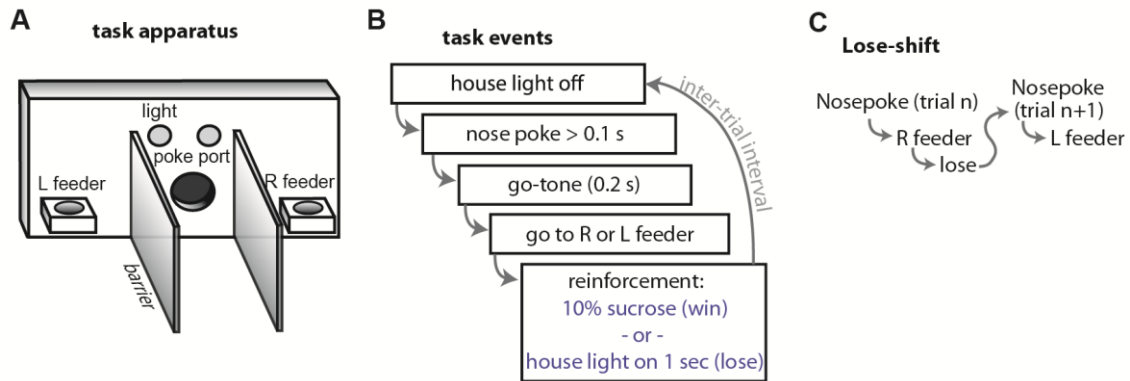


Figure 1. Matching Pennies schematic diagram and flow chart detailing a trial and lose-shift responding. (A) The operant chamber the task is performed in. (B) The sequence of events that encompass a trial in the MP task. (C) Schematic detailing an example of lose-shift responding.

Recent work from the Gruber lab has implicated the dorsolateral sub-region of striatum as being the locus of lose-shift responding. Lesions to dorsolateral striatum (DLS) reduced lose-shift responding in the MP task, suggesting that it plays a role in promoting choice shifts following undesirable choice outcomes. This idea is supported by its innervation from dopaminergic substantia nigra neurons (Haber et al., 2000; Joel and Weiner, 2000), which can transmit RPE signals. The other sub-regions of striatum also receive afferent projections from dopamine neurons. Specifically, VTA dopamine neurons project to the dorsomedial and ventral sub-regions of striatum (Joel and Weiner, 2000). Although the boundaries of these sub-regions are more akin to gradients than distinct lines (Voorn et al., 2004), each sub-region has been assigned functionally segregated roles in decision-making.

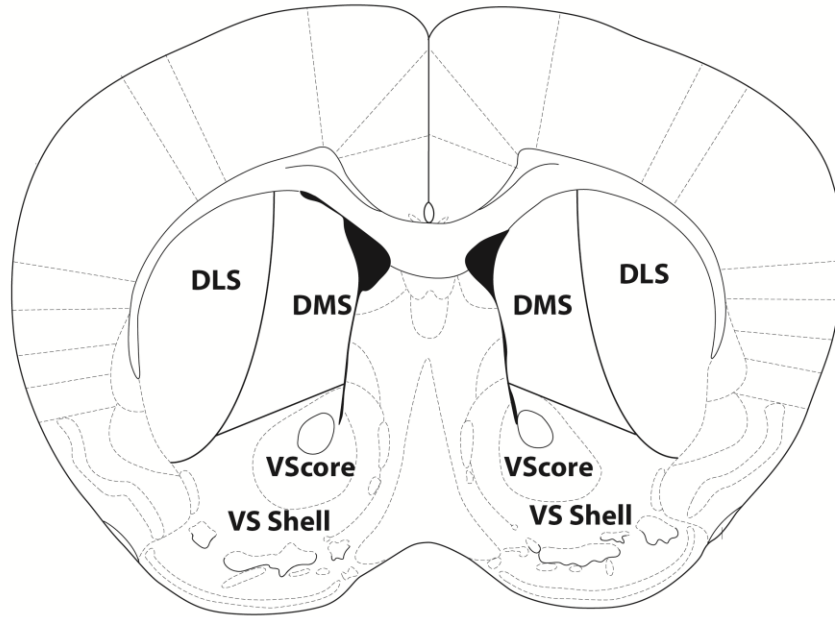


Figure 2: Coronal section of rat brain detailing striatal sub-regions.

Dorsomedial striatum (DMS) has been proposed to play a role in ‘goal directed’ decision-making (Yin et al., 2005b). Ventral striatum (VS) appears to modulate response vigor (Caul and Brindle, 2001) and the motivation to engage or disengage in tasks, as determined by the expected value of return (Nicola, 2010). Finally, dorsolateral striatum (DLS) is proposed to learn stimulus-response associations that can form ‘habits’. This is distinct from forming associations with outcome values, and is consistent with studies showing that rats with lesions to this region are insensitive to outcome devaluation (Yin et al., 2004). While these segregated roles have formed the predominant view of striatal influences on decision-making, there is a growing body of research that suggests these assigned roles are incomplete. For example, restoring dopaminergic function to DLS in dopamine-deficient mice increases performance on instrumental tasks, conditional discrimination, and other ‘goal-directed’ cognitive behaviors (Darvas and Palmiter, 2009). Furthermore, rats with DLS lesions (and presumably greater DMS control over

behavior) have been found to utilize egocentric navigation, which would normally be attributed to DLS (Chang and Gold, 2004). A final example is DLS being involved in the reward-sensitive, short timescale, lose-shift responding, which is in direct opposition to the idea that the region is the slow-learning, reward-insensitive, ‘habit’ system. These findings would not be expected by predominant theory and suggests our knowledge of striatal influence on decision-making is incomplete.

1.3 Drug sensitization, synaptic plasticity, and decision-making

Repeated drug administration causes well-documented changes in the morphology of the medium spiny neuron (MSN), which is the predominant cell type in striatum (Tepper and Bolam, 2004). AMPH sensitization increases dendritic spine density, branching, and length in VS (Robinson and Kolb, 1997), as does cocaine (Robinson and Kolb, 1999). Furthermore, methamphetamine sensitization causes increased spine density in DLS and VS, while decreasing spine density in DMS (Jedynak et al., 2007).

These plastic changes have been proposed to facilitate the progression of addiction by increasing the relative behavioral control of DLS over other sub-regions (Willuhn et al., 2012; Everitt and Robbins, 2013), leading to habitual drug use without regard for potential negative consequences. The idea of competing behavior systems in striatum is illustrated by a study demonstrating that lesions of the ‘goal-directed’ DMS (which presumably increase behavioral control by the intact DLS) leave animals insensitive to reward devaluation, as they have become habitual responders to a stimulus (Yin et al., 2005b). An analogous process may occur in addiction with drug-induced plastic changes promoting drug use as a habit rather than something that is desired.

Given the well-documented plastic changes in striatum from drug sensitization, we hypothesize that DLS will have relatively greater control over behavior in our MP task in sensitized animals. We predict this will manifest as increased lose-shift responding during the task. In contrast to this, we expect that acute administration of AMPH and THC will reduce lose-shift responding by attenuating the negative prediction error normally associated with losses. (see Chapter 5. General Discussion for a direct comparison of the behavioral effects of AMPH and THC).

2. Opposing effects of acute and chronic d-amphetamine on decision-making in rats

Abstract:

Amphetamine and other drugs of abuse have both short-term and long-lasting effects on brain function, and drug sensitization paradigms often result in chronic impairments in decision-making. Here we show that acute amphetamine administration temporarily renders rats less sensitive to reward omission, as revealed by a decrease in lose-shift responding during a binary choice task. Intracerebral infusions of amphetamine into VS did not affect lose-shift responding but did increase impulsive behavior in which rats chose to check both reward feeders before beginning the next trial. In contrast to acute systemic and intracerebral infusions, sensitization through repeated exposure induced long-lasting increased sensitivity to reward omission. These treatments did not affect choices on trials following reward delivery (i.e. win-stay responding), and sensitization increased dendritic spine density in dorsolateral striatum. The dichotomous effects of amphetamine on short-term and long-term loss sensitivity, and the null effect on win-stay responding, are consistent with a shift of behavioral control to DLS after drug sensitization. These data provide a new demonstration of such a shift in a novel task unrelated to drug administration, and suggests that the dominance of sensorimotor control persists over many hundreds of trials after sensitization.

2.1 Introduction

Animals learn from reinforcement by a neural mechanism thought to involve dopamine. Dopaminergic neurons in the midbrain code for a reward prediction error signal by increasing firing rate when an unexpectedly good reward is presented, and decrease firing when a smaller than expected reward or the omission of a reward occurs (Schultz et al., 1997; Fiorillo et al., 2003; Roesch et al., 2007). This error signal is the basis for many models describing how humans and animals can use trial-and-error learning to make beneficial decisions in novel environments or tasks (Montague et al., 1996; Frank et al., 2004; Pessiglione et al., 2006).

Dopamine neurons densely innervate striatum, a structure strongly implicated in reinforcement-based learning (O'Doherty et al., 2006; Frank et al., 2007; Johnson et al., 2007; Ito and Doya, 2009; Kimchi and Laubach, 2009). The rodent striatum is often conceptually divided into ventral, dorsomedial, and dorsolateral sub-regions, which are thought to be homologous to nucleus accumbens, caudate, and putamen in primates (Balleine and O'Doherty, 2010). It has been suggested that these sub-regions are components of parallel circuits between the cortex, basal ganglia, and thalamus (Alexander et al., 1986; Haber, 2003; Voorn et al., 2004). These circuits appear to have distinct information processing capabilities and can interact to control decision-making. For instance, instrumental conditioning paradigms have suggested that DMS encodes action-association outcomes (Yin et al., 2005a) that allow it to form mental models of its environment (Daw et al., 2005). In contrast, DLS appears to encode stimulus-response associations that are built up over repetition without such models (Packard and McGaugh, 1996; Jog et al., 1999; Featherstone and McDonald, 2004) and are insensitive to altered

reward contingencies such as devaluation (Yin et al., 2004). These are generally conceptualized as habits that are engaged reflexively (Jog et al., 1999).

We recently discovered that DLS mediates so-called lose-shift (or lose-switch) responding, wherein animals tend to shift responses to an alternate option following reward omission (Skelin et al., 2014). This is important for two reasons: this strategy may influence behavioral flexibility, particularly after reward contingencies change; and it is a new behavioral barometer of the prevalence of DLS-mediated control of choice. DLS thus influences behavioral flexibility in normal and drug-induced states by affecting animals' responses after reward omission by promoting lose-shift responding. Acute (on board) amphetamine increases extracellular dopamine, norepinephrine, and other monoamines, particularly in the uptake-transporter rich striatum (Pontieri et al., 1995; Heien et al., 2005). We hypothesized that acute amphetamine (AMPH) would decrease lose-shift by attenuating the negative RPE signal associated with a loss. On the other hand, sensitization has been posited to shift the control of behavior to DLS (Everitt and Robbins, 2005; 2013; Lucantonio et al., 2014). This would lead to the dominance of DLS-driven responses in behavioral control. As such, we expect an increase in lose-shift responding following AMPH sensitization. This shift of control to sensorimotor systems may correspond with the increase in the dendritic spine density in DLS, and decreases in DMS and orbitofrontal cortex (OFC), induced by sensitization (Crombag et al., 2005; Jedynak et al., 2007). Our findings support these hypotheses, and highlight lose-shift responding as a novel measure of DLS function in models of addiction.

2.2 Methods

2.2.1 Subjects

Subjects were 32 adult male Long-Evans rats (Charles River, Quebec) weighing 250-350 g. Animals were pair-housed in a climate-controlled vivarium under a 12:12 hour light:dark cycle (lights on 7:30 a.m.). Animals were given access to water for one hour on behavioral testing days, but otherwise had ad libitum access to food and water. All procedures were approved by the University of Lethbridge Animal Welfare Committee, following the guidelines of the Canadian Council on Animal Care.

2.2.2 Apparatus and choice task

Behavioral testing was performed in aluminum operant chambers (26 x 26 cm) containing two cue lights and a central port flanked by two sucrose water delivery feeders (see Skelin et al. (2014) for details). The central port and sucrose water delivery feeders contained infrared sensors to detect entry and exit. For behavioral testing, animals were placed in the operant chamber for one hour sessions. Control of the task was automated by an Arduino Mega microcontroller (Arduino, Italy) receiving commands via custom software on a computer. Illumination of the cue lights indicated the beginning of a new trial, signaling the animal to nose-poke in the central port. A tone (6 kHz; 150 ms duration) then prompted the animal to select one of the two sucrose delivery feeders. If the correct feeder was chosen, a reward (60 μ L of 10% sucrose solution) was delivered after a 500 ms delay. If the incorrect feeder was chosen no sucrose was delivered. Once a feeder was chosen, or if no feeder was chosen in the 15 s following a nose-poke, the trial ended and the animal had to return to the central port to initiate a new trial.

In the first session of behavior shaping, animals were rewarded upon every feeder entry following a nose-poke in the central port to train them to perform the nose-poke and feeder entry sequence. In the second session, the probability of reward was 50% for each feeder entry following a nose-poke to teach them that not all valid responses lead to reinforcement. In all subsequent sessions, reinforcement was controlled by an algorithm that attempted to minimize the number of rewards given to the animal by predicting which feeder it would select. This was done by examining the choices and reinforcements from the previous four trials (Lee et al., 2004). If either feeder was selected at a greater than chance rate (probability > 0.5 with the binomial test, $P < 0.05$), it would be unrewarded for the upcoming trial. The task thus implements a two-player competitive choice task, which is sometimes called ‘Matching Pennies’. Over consecutive days of training, two small (4.0 cm), medium (8.5 cm), or long (13.5 cm) parallel barriers were added to the operant chamber to separate the central nose-poke port and the feeders. This introduced a choice cost by forcing animals to navigate around it and also reduced feeder bias from body orientation by promoting posture that was orthogonal to the wall in which the port and feeders were mounted. Rats were trained until they completed two consecutive sessions of at least 150 trials with the long barriers. This criterion was met by training session 13 for all rats in the study. All subsequent training and testing sessions were run with the long barriers.

2.2.3 Drug preparation and injections

For Experiments 1 and 2, d-amphetamine hemisulfate (Sigma-Aldrich, Ontario) was dissolved in 0.9% sterile saline at three different concentrations so that animals received approximately the same injection volume across dosages. AMPH solution was

delivered by intraperitoneal (IP) injection at one of three dosages (0.5mg/kg, 1.0mg/kg, or 1.5mg/kg) for Experiment 1, and an escalating dose (1.0mg/kg, 2.0mg/kg, and 2.0mg/kg twice per day) was given in Experiment 2. Injection sites were rotated and alternated sides to minimize irritation. For Experiment 3, AMPH (0, 20, and 40 $\mu\text{g}/\mu\text{l}$) was made in the same manner as the previous experiments using Phosphate-Buffered Saline (PBS) rather than sterile saline, and was infused into DLS or VS.

2.2.4 Experiment 1: Acute effects of AMPH on the MP task

After initial shaping, animals were randomly divided into four groups of four to receive acute AMPH in a counterbalanced block design. Injections were administered 15 minutes prior to testing on the behavioral task over a period of 8 days using the following schedule: saline, injection 1, no injection, injection 2, no injection, injection 3, no injection, and injection 4. The initial saline injection was to habituate animals to the procedure and was not used for analysis. Injection days consisted of one of the three AMPH dosages or vehicle (saline). The order in which the drug dosages were administered on injection days was counterbalanced across subjects.

2.2.5 Experiment 2: Effects of AMPH sensitization on the MP task

Thirty days after the conclusion of the acute AMPH experiment, animals were divided into amphetamine ($n = 8$) and control ($n = 8$) groups and subjected to three weeks (five days on, two days off) of intraperitoneal injections of AMPH. The drug schedule followed an escalating dose (week 1: 1mg/kg once a day; week 2: 2mg/kg once a day; week 3: 2mg/kg twice a day) similar to previous studies (Robinson and Kolb, 1997).

After a ten day drug abstinence period, animals received three retraining days followed by three days of behavioral testing with the competitive mode (one session per day).

2.2.6 Experiment 3: Effects of AMPH infusion into DLS and VS on the MP task

2.2.6.1 Surgery

A separate cohort of rats were divided into DLS (n=8) or VS (n=8) groups for bilateral cannulae implantation. Animals were first given subcutaneous injections (0.05mg/kg) of Buprenorphine (Alstoe Ltd., UK) 30 minutes before making the first incision, and were anesthetized using 4% isoflurane gas (Benson Medical Industries Inc., Ontario) in oxygen flowing at 1.0 L/min. Surgical plane was maintained using 2% isoflurane throughout the surgery. Two burr holes were drilled into the skull for the bilateral implantation of 8 mm 23-gauge stainless steel guide cannulae (Plastics One, VA) into either DLS or VS at the following coordinates from bregma [in mm (AP, ML, DV)]: DLS (0.9, 3.6, -4.3); VS (1.56, 2.8, -7.2), with VS cannulae angled 10 degrees toward the midline to avoid hitting the central sinus. Three anchoring screws were inserted into the skull, and the cannulae were held in place using dental acrylic. The guide cannulae were occluded using an 8mm 30-gauge infusion dummy cannula that fit flush with the guide cannula. Following surgery, rats were given subcutaneous injections (0.02mg/kg) of meloxicam (Boehringer Ingelheim, Germany) and monitored for 24 hrs before returning them to their home cages. The animals recovered in their home cages (pair housed) for one week before resuming training.

2.2.6.2 Behavior testing

Rats had three post-surgery re-training days on the task before infusions began. Fifteen minutes before behavioral testing, rats were gently restrained as the dummy cannulae were removed and the infusion internal cannulae were inserted. A total of 0.5 μ l of AMPH solution (20, and 40 μ g/ μ l) or PBS was infused into each site at 0.30 μ l/min with an infusion pump (Harvard Apparatus, Massachusetts). The internal infusion cannulae were left in place for three minutes post-infusion to allow for diffusion of drug. All rats were infused with saline on the first day of testing to habituate them to the infusion procedure and to clear any clogs in the guide cannulae. These days were not used for analysis. Each animal received the two doses of AMPH and PBS on three separate testing days in a counterbalanced block design with testing on infusion days alternating with non-infusion days (as in Experiment 1).

2.2.6.3 Histology

Upon completion of behavioral testing, the animals received lethal injections of sodium pentobarbital (100 mg/kg IP) and were perfused with PBS and 4% paraformaldehyde (PFA). The brains were post-fixed for 24 hours in 4% PFA and then cryoprotected in 30% sucrose in PBS solution with sodium azide (0.02%). Brains were then sectioned in the coronal plane at 40 μ m thickness using a SM2010R freezing microtome (Leica, Germany). Every fourth section of striatum was wet-mounted on glass microscope slides and labeled with cresyl violet. Sections were photographed using a NanoZoomer (Hamamatsu, Japan).

2.2.7 Behavior analysis

Data were analyzed using Matlab version 2013a (Mathworks, Massachusetts) and Graphpad Prism version 6 (Graphpad, California). Rats sampled both reward feeders on some trials, a phenomenon we refer to here as impulsive feeder approach (IFA). We omitted the subsequent trials from analysis of lose-shift and win-stay probabilities in order to minimize potential confounds introduced by visiting both feeders. For Experiment 1 (Acute AMPH), we examined several features of responding including: lose-shift, win-stay, reaction time, number of trials, response entropy, and IFA using one-way repeated measures Analysis of Variance (ANOVA) tests. Lose-shift responding is a measure of loss sensitivity wherein animals chose to shift feeder choice from that of the previous trial if they failed to receive reinforcement on that trial. Win-stay is the re-selection of the feeder that provided a reward on the previous trial. Response entropy is a measure of the randomness of an animal's choice sequence (Skelin et al., 2014). Impulsive feeder approach is the out of trial sampling of feeders for reward. Unpaired t-tests were used to test differences of group-averaged means (sensitized vs. non-sensitized) in Experiment 2. Two-way mixed-model repeated measures ANOVAs were used to analyze the effect of drug dose and intracerebral infusion (Experiment 3).

2.2.8 Tissue/diolistic bullet preparation and diolistic labeling

DiI (1,1'-Dioctadecyl-3,3',3'-Tetramethylindocarbocyanine perchlorate)-coated microbeads were created using methods described previously (Seabold et al., 2010). Briefly, 1.3 μm tungsten microbeads (Bio-Rad, California) were coated in lipophilic DiI and dried on the inside of plastic tubing (Saint-Gobain, France). Once dry, tubing was cut

into 13mm ‘bullets’ that were attached to the nozzle of a home-made ‘gene gun’ comprised of a solenoid valve that was electronically briefly opened to allow compressed air to pass through. We used a bolus (50 ms valve opening) of compressed air (60 psi) to expel the beads into tissue at a distance of 8 cm. Each section of tissue was labeled once to prevent overlabeling. Tissue was prepared following behavior testing by transcardial perfusion of 150ml of PBS, followed by 150ml 4% PFA. Brains were then post-fixed for two hours at room temperature in 4% PFA. Striatal sections were cut at 200 μm using a Leica VT1000 S vibratome after stabilizing the brain with an agarose block to improve cutting quality. Sections were individually placed in separate chambers of a 24 well plate and minimally submerged under PBS to keep them moist. The PBS was removed by pipette prior to diolistic labeling and then immediately covered with PBS after labeling. Sections were then washed three times in PBS and incubated overnight in the dark, while shaking gently at room temperature. The sections were then fixed again with 4% PFA for 2 hours, washed with PBS 3 times, and mounted and coverslipped using Vectashield with Dapi (Vector Labs, California).

2.2.9 Confocal imaging and dendritic spine density analysis of DiI-impregnated tissue sections

A Nikon TE2000-U confocal microscope was used to image labeled tissue sections. DiI was excited at 543nm using a Helium-Neon laser (Reo Inc., CO). Striatal MSNs were first identified at low magnification, and sections of dendrite from that neuron were imaged using a 60x water-immersion lens (Plan-Apo, Nikon, NA = 1.2, WD = 270 μm) with an optical zoom of 4.5. Distinct image stacks of dendrites were taken as distal as possible from the cell body, with sequential stacks moving closer to the soma up

to but not passing the first branch point. Only dendrites that were clearly connected to cell somas and were distinct from other dendrites were imaged for analysis. Confocal stacks were taken with a frame size of 512x512 pixels, a 0.15 μ m step size in the z plane up to 250 steps, and each individual step was averaged over three images to reduce background noise. Image stacks were then deconvolved using the adaptive 3-D blind deconvolution feature in AutoQuant X version 3.02 (Media Cybernetics, Maryland), similar to the procedure by Jung et al. (2013). These image stacks were then loaded into Imaris version 7.6.5 (Bitplane, Connecticut) for semi-automated dendritic spine analysis.

Spine density analysis was done using similar methods to those described previously (Shen et al., 2008; Shen et al., 2009). Briefly, the Surpass module of Imaris was used to reconstruct a three-dimensional image. Dendritic spine maximum length and diameter were then measured using the Slice module. In the Filament module, dendrites were manually traced and the dendrite diameter manually adjusted to match that of the image stack template. Spines were then modeled using the spine detection wizard using spine seed point diameter and spine maximum length as previously measured.

2.3 Results

2.3.1 Experiment 1 (acute effects of AMPH on the MP task)

We trained 16 rats to perform a binary choice task with no explicit cues as to the reward outcome at either of two feeders (see Methods). Although random feeder selection from trial to trial is the optimal policy on this task, we previously found that rats persistently used a lose-shift strategy (Skelin et al., 2014). We first tested how acute administration of AMPH affects choice behavior on the task (see Fig. 1). Reaction time significantly decreased ($F(3,45)=7.63$, $p<0.01$; Fig 1.A) in a dose-dependent manner with

increasing AMPH, consistent with previous reports of motoric increases by this drug (Wilkinson et al., 1993). However, AMPH did not significantly increase the number of trials per session ($F(3,45)=1.61, p=0.21$). We found that rats intermittently sampled both reward feeders during the inter-trial interval, and this was increased by AMPH ($F(3,45)=11.5, p<0.001$; Fig 1.B). This sampling of both feeders may affect choice. We therefore limited the analysis of reinforcement on choice to trials with intertrial intervals of less than 5 seconds, which eliminates the potential effect of sampling both feeders and other behaviors (e.g. grooming) that may disengage animals from the task and/or interfere with the representation of reinforcement information between trials. Lose-shift responding significantly decreased in a dose-dependent manner ($F(3,45)=6.44, p<0.01$; Fig 1.C). No effect of AMPH was found on the probability of win-stay responding ($F(3,45)=2.01, p=0.15$; Fig 1.D). Lastly, AMPH did not affect response entropy ($F(3,45)=2.16, p=0.13$), suggesting that it is the sensitivity of choice on reinforcement that is affected by the drug rather than an overall change in response pattern.

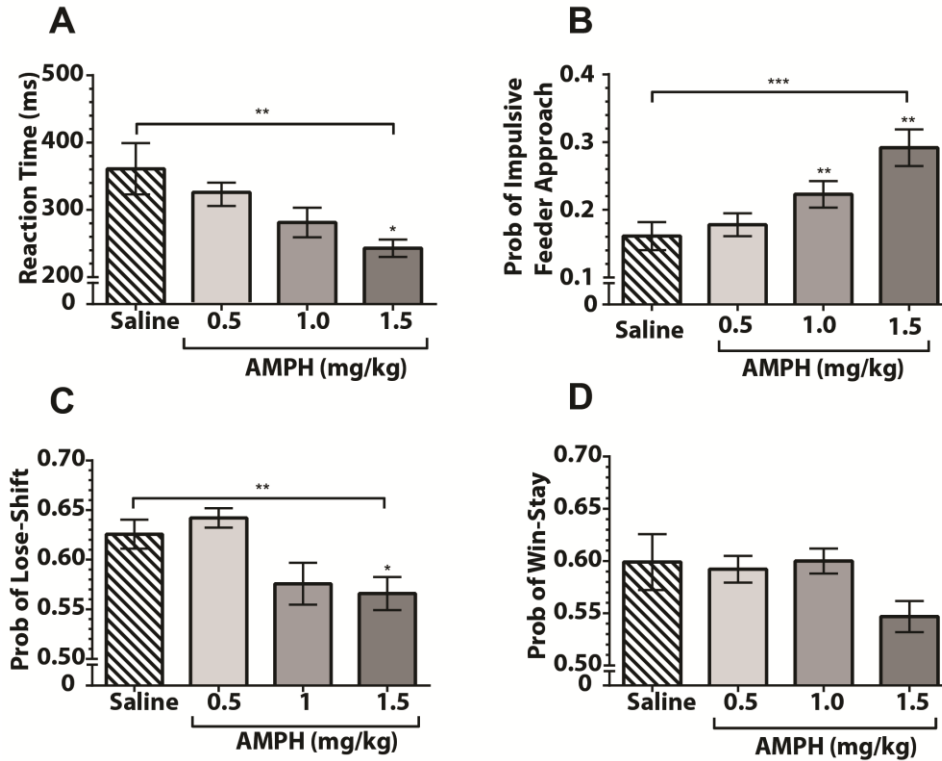


Figure 3. Effects of acute AMPH administration on behavioral measures. (A) Group-averaged reaction time, showing dose-dependent motoric speeding. (B) Probability that animals sampled both reward feeders between trials, which was increased by AMPH. (C) Probability of lose-shift responding, which was decreased by AMPH. (D) Probability of win-stay responding was not affected by AMPH. A significant main effect of drug dose by ANOVA is indicated by ‘†’, and asterisks above individual columns indicates group means that were significantly different from the control (Saline) mean according to Dunnett's post hoc test with $\alpha = 0.05$.

2.3.2 Experiment 2 (Effects of AMPH sensitization on the MP task)

We next investigated if sensitization by repeated AMPH administration affects choice behavior in the choice task. The rats from Experiment 1 were divided into two groups. One group (n=8) received escalating daily injections of AMPH over 3 weeks, and the other group (n=8) received vehicle injections (saline) following the same schedule. A 10-day abstinence period followed the end of the injection sequence, and the rats were then tested on the task. Responding by each rat was averaged over the three testing

sessions prior to computing means for the drug and control groups, and the significance of the difference of these group means was compared using unpaired t-tests. No effect was seen on the number of trials per session ($t(14)=0.024$, $p=0.49$) or reaction times ($t(14)=0.529$, $p=0.30$; Fig 1.A) in the groups. Unlike acute AMPH, sensitization had no significant effect ($t(14)=1.51$, $p=0.15$; Fig 1.B) on impulsive feeder approach between AMPH-sensitized and Saline groups, although a trend of increased sampling exists across all three days of testing. In contrast to acute drug, amphetamine sensitization significantly *increased* lose-shift responding ($t(14)=2.59$, $p=0.02$; Fig 1.C) between AMPH-sensitized and Saline groups. Similar to acute AMPH administration, no effect was found on win-stay ($t(14)=0.177$, $p=0.86$; Fig 1.D) or response entropy ($t(14)=1.147$, $p=0.14$).

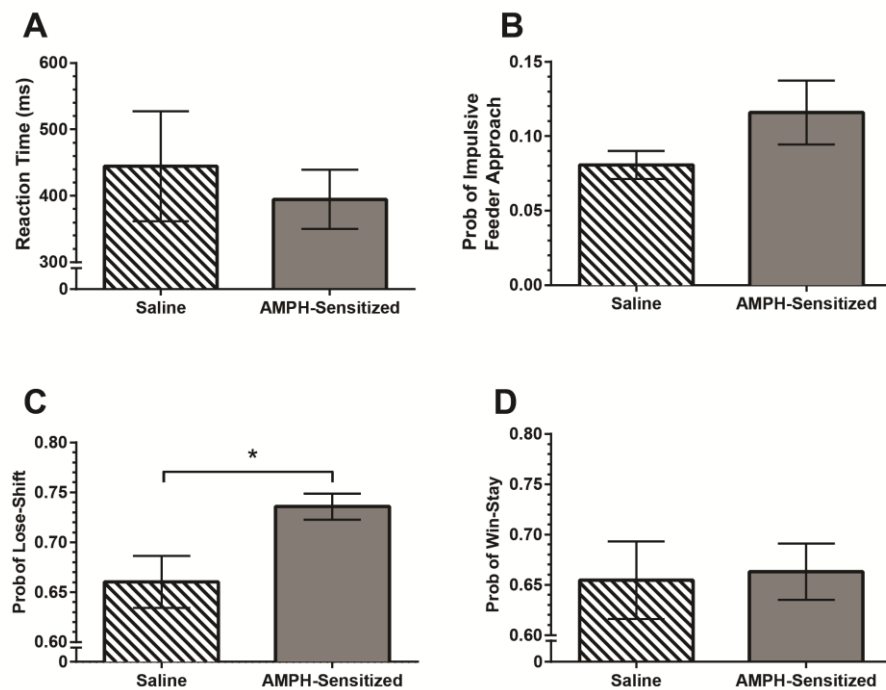


Figure 4. Effects of AMPH sensitization on behavioral measures. (A) Sensitization had no significant effect on reaction time. (B) Impulsive feeder approach was unaffected by sensitization. (C) AMPH sensitization increased the probability of lose-shift responding. (D) Win-stay was not affected by drug sensitization. Significant differences ($p<0.05$) between groups are denoted by asterisks.

2.3.3 Dendritic spine density analysis

We next sought to test if the AMPH sensitization increased dendritic spine density in dorsolateral striatum. We used ‘diolistic’ labeling to impregnate neural membranes with a fluorescent lipophilic dye (DiI), and used confocal imaging to then obtain a 3-dimensional stack of the fluorescent neuron. Twenty-nine neurons from ten animals (5 control; 5 sensitized) used in Experiments 1 and 2 were analyzed by an experimenter without knowledge of the experimental conditions. When multiple neurons or dendrites were analyzed from the same animal, they were averaged to get a single spine density measure for each animal. Analysis showed that the AMPH-sensitized animals had significantly greater spine density in DLS ($t(8)=4.932$, $p=0.001$; Fig 3.E).

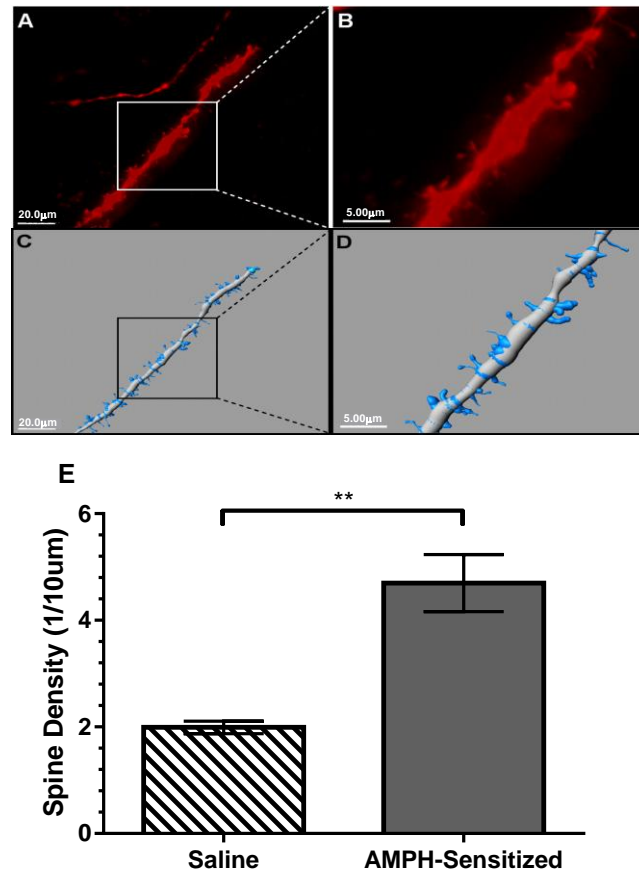


Figure 5. Spine density quantification in DLS. (A) Representative sample of a deconvolved and reconstructed confocal image stack. (B) Enlarged section of panel A. (C) Morphological model of the dendrite in panel A. (D) Enlarged model corresponding to panel B. (E) Group averaged spine density, showing significantly more spines in the dorsolateral striatum of AMPH-sensitized animals. Significant differences between groups are denoted by asterisks.

2.3.4 Experiment 3: Acute effects of AMPH intracerebral infusion on the MP task

We have shown here that distinct behaviors are affected differently by systemic AMPH administration; specifically lose-shift responding decreased whereas the approach to both feeders increased. We next sought to determine if these dissociated behavioral effects of systemic AMPH derive from the action of the drug in distinct regions of striatum. To investigate the possible involvement of distinct behavioral effects of AMPH in VS versus DLS in the task, we trained 16 animals on the task and then bilaterally implanted infusion cannulae into the VS (n=8) or DLS (n=8) to deliver drug directly into these target regions (Fig 4.A/B). We infused either saline, low dose AMPH (20 μ g/ μ l), or high dose AMPH (40 μ g/ μ l) 15 minutes before testing in a counterbalanced block design. There was no effect of drug ($F(2, 28) = 1.246, p=0.30$) or brain region ($F(1, 14) = 4.182, p=0.06$) on reaction time (Fig 4.C), although brain region approached significance. However, there was a significant increase in the number of trials with drug ($F(2, 28) = 8.467, p<0.01$), while brain region had no effect ($F(1, 14) = 0.02333, p=0.88$). There was also a significant effect of drug ($F(2, 28) = 3.708, p=0.04$, and brain region ($F(1, 14) = 11.31, p<0.005$) on impulsive feeder approach (Fig 4.D), with the drug increasing this behavior in VS but not DLS. Lose-shift responding (Fig 4.E) was not significantly affected by drug dose ($F(2, 28) = 1.495, p=0.24$) or brain region ($F(1, 14) = 0.6912,$

p=0.42). There was also no main effect of drug ($F(2, 28) = 1.002, p=0.38$) or brain region ($F(1, 14) = 3.774, p=0.07$) on win-stay responding (Fig 4.F), although brain region had a strong trend such that win-stay decreased with increasing AMPH in VS but not DLS. This is consistent with the trend of decreasing win-stay with acute systemic administration. Finally, response entropy was not significantly affected by drug dose ($F(2,28) = 1.69, p=0.70$) or brain region ($F(1,14) = 0.02, p=0.89$).

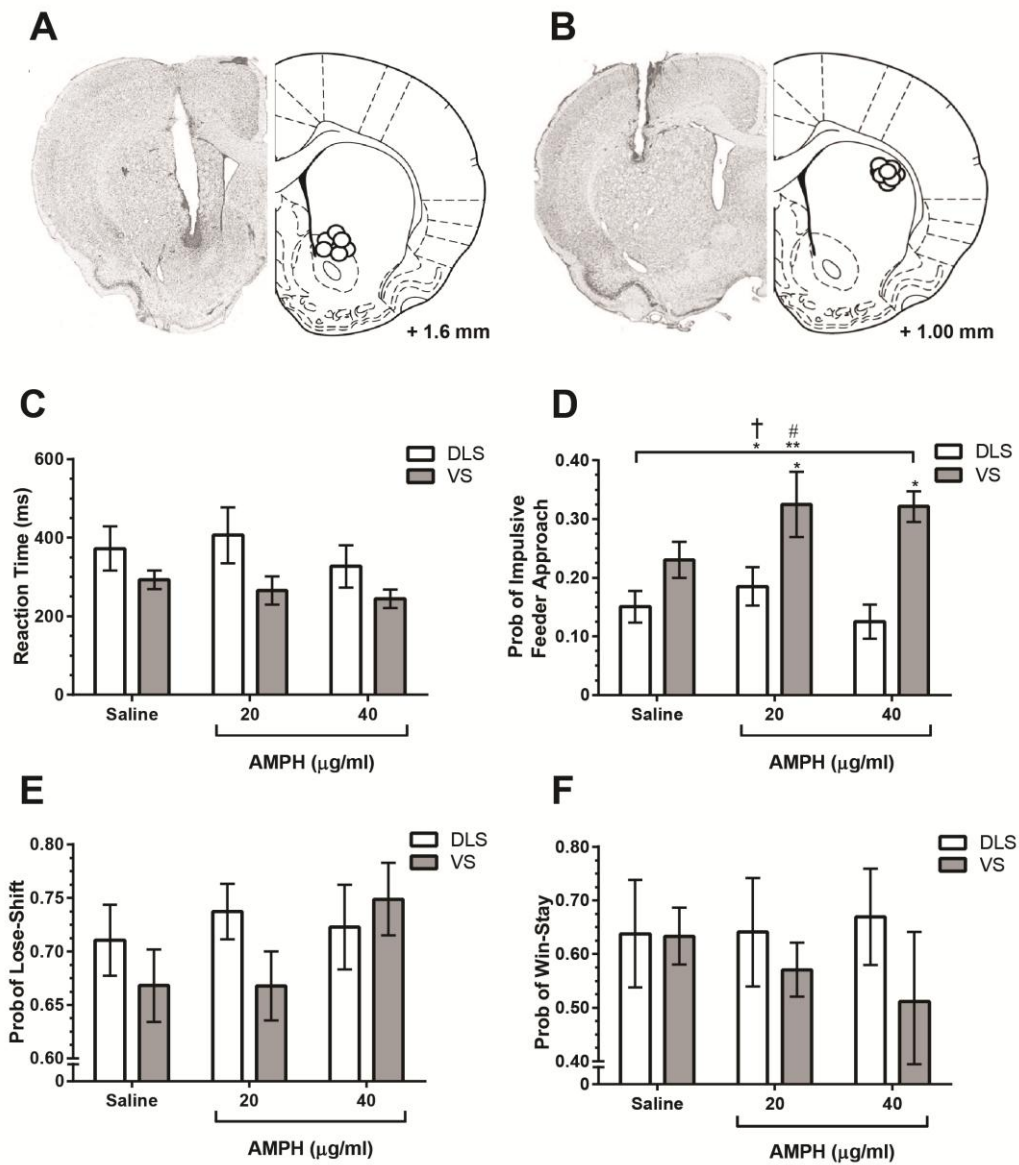


Figure 6. Histology and behavioral effects of AMPH infusion into DLS or VS. Schematic diagram and microphotography of coronal sections showing the area of microinfusion site in (A) VS and (B) DLS. Cannulae were placed bilaterally. The black circles represent approximate infusion site in each animal. (C) Infusion had no significant effects on reaction time. (D) Drug dose and brain region had significant main effects on the probability that animals sampled both reward feeders between trials. (E) Probability of lose-shift responding, which was unaffected by AMPH. (F) The probability of win-stay responding was not affected by AMPH. A significant main effect of drug dose is indicated by ‘†’, while a significant main effect of brain region is indicated by ‘#’. Asterisks above individual columns indicates group means that were significantly different from their control (saline) mean according to Dunnett's post hoc test with $\alpha = 0.05$.

Table 1. Group averages of number of trials per session and response entropy for AMPH experiments. Standard deviation indicated in brackets.

	Acute Amphetamine				Sensitization		Infusions					
	Saline	0.5 mg/kg	1.0 mg/kg	1.5 mg/kg	Saline	Sensitized	DLS		VS		VS	
							PBS	20µg/ml	40µg/ml	PBS	20µg/ml	40µg/ml
Number of Trials/Session	410.6 (61.5)	422.4 (66.3)	448.7 (46.9)	427.9 (61.0)	288.9 (30.6)	289.5 (55.9)	307.4 (63.1)	337.6 (60.9)	356.5 (67.9)	311.4 (45.3)	308.1 (49.0)	392.6 (40.6)
Response Entropy (bits)	3.845 (0.08)	3.873 (0.05)	3.856 (0.09)	3.897 (0.06)	3.819 (0.05)	3.750 (0.15)	3.835 (0.09)	3.798 (0.13)	3.854 (0.04)	3.856 (0.07)	3.856 (0.05)	3.788 (0.16)

2.4 Discussion

The goal of this study was to exploit lose-shift responding and impulsive feeder approach as novel behavioral assays to test hypotheses about the effects of acute and repeated psychostimulant administration on reinforcement-driven choice adaptation. We found that acute AMPH *decreased* lose-shift responding rates in a dose-dependent manner, whereas sensitization *increased* lose-shift responding rates. The former finding suggests that lose-shift responding involves dopamine or other neuromodulators affected by AMPH, and the latter suggests that the brain mechanisms supporting lose-shift become more dominant in the control of behavior after sensitization. Because we have previously shown that lose-shift responding is disrupted most profoundly by lesions of DLS (Skelin et al., 2014), our data provides novel support for the theory that models of drug abuse cause a shift of behavioral control to the sensorimotor systems involving DLS (Everitt

and Robbins, 2013). We also found here that the propensity of rats to sample both feeders between trials (IFA) was increased by acute systemic AMPH, or by infusions directly into VS but not DLS. These data suggest that lose-shift responding and IFA involve dissociated striatal circuits that are differently affected by AMPH. Lastly, win-stay responding was not significantly affected by any drug manipulations here, but did show a trend of decreasing with AMPH in VS. This could be a third dissociated behavior, but the present evidence is weak due to the possibility of a false negative.

On-board AMPH has multifaceted effects in the brain and increases extracellular dopamine by reversing DAT function (Fleckenstein et al., 2007) and serves as a substrate for dopamine exchange with the vesicular monoamine transporter (Eiden and Weihe, 2011). This supraphysiological dopamine release is believed to attenuate the effect of pauses in dopamine neuron firing triggered by worse-than-expected outcomes, such as reward omission in the present task (Schultz, 1998; Nakahara et al., 2004; Pan et al., 2005). The reduction of lose-shift responding by systemic AMPH is consistent with the notion that low dopamine states promote response shifts. Acute injections of systemic AMPH also impair reward-based decision-making in other tasks, such as a modified rat Iowa Gambling Task (Zeeb et al., 2009) and reversal learning task (Idris et al., 2005). These results may also be explained by reduced sensitivity to poor reward outcomes and/or impaired lose-shift responding. Our previous findings that lose-shift depends most strongly on DLS suggest that it is the dopamine neurons from substantia nigra, which selectively innervate DLS (Haber et al., 2000; Joel and Weiner, 2000), that are involved in triggering lose-shift. Indeed, recent work has shown a similar functional role of dopamine neurons in substantia nigra and ventral tegmental area (Ilango et al., 2014).

However, local infusion of AMPH into DLS through a single cannula in each hemisphere did not significantly decrease lose-shift. We speculate that this could be a false negative due to a limited spatial extent of drug infusion, effects of motoric speeding, or other confound. In further support of this idea, we have published an abstract reporting that infusion of an agonist for D2-type dopamine receptors through two bilateral cannulae in DLS, but not in DMS, reduces lose-shift responding (Thapa et al., 2014). Thus, we suggest the infusion data here are not strong evidence against a role of DLS in lose-shift responding.

We also cannot rule out the involvement of other neurotransmitters.

Amphetamine facilitates the release of all catecholamines and serotonin (Sulzer et al., 2005), presenting a confound. For example, high doses of the selective serotonin reuptake inhibitor citalopram reduces loss sensitivity on a reversal learning task (Bari et al., 2010). Inhibition of norepinephrine reuptake also decreases the number of trials needed to reverse choice discrimination during reversal learning in rats and monkeys (Seu et al., 2009). To properly determine whether other neurotransmitter systems play a role in our task it will be necessary to use more selective manipulations in future experiments.

In contrast to the reduced lose-shift responding induced by acute AMPH, a sensitization regime increased lose-shift responding. The most likely explanation derives from previous proposals that such sensitization causes a shift in behavioral control towards DLS (Willuhn et al., 2012; Everitt and Robbins, 2013), which is supported by electrophysiological data (Thorn et al., 2010). Here we used sensitization as a simple model of addiction, and test how this affects decisions in a non-drug context. Our results suggest that decision-making in general is shifted to DLS control. Physiologically, this

transition of control may be produced by plastic changes in striatum. Psychostimulant drugs exert strong effects on plasticity in the brain (Robinson and Kolb, 2004).

Amphetamine sensitization *increases* dendritic spine density, dendrite length, and branching in VS, medial prefrontal cortex (Robinson and Kolb, 1997; 1999), and DLS (Jedynak et al., 2007), while it *decreases* spine density in DMS (Jedynak et al., 2007). Consistent with this, we find greatly increased DLS spine density in amphetamine-sensitized rats. It is worth noting that our calculated spine density is lower than other reported values. This may be due our use of 4% PFA rather than 1.5% PFA; as higher fixative concentration has been proposed to cause underlabeling (Staffend and Meisel, 2011). The greater spine density in DLS and VS as compared to DMS may thus lead to a shift in behavioral control to these structures. Indeed, lesions of DMS produce deficits in reversal learning (Clarke et al., 2008; Castañé et al., 2010) and set shifting (Lindgren et al., 2013), tasks that are also impaired by psychostimulant sensitization (Jentsch et al., 2002; Schoenbaum et al., 2004; Featherstone et al., 2008).

We argue that our task is also sensitive to VS-driven behavioral control, as represented by impulsive feeder approach. Acute AMPH had a strong dose-dependent effect on the propensity of animals to sample both feeders between trials, and a similar trend (although not significant) was present during every day of testing after sensitization. Furthermore, infusions of AMPH into VS (but not DLS) also caused significant increases in impulsive feeder approach. Such IFA was never reinforced at any time during shaping or training, and so it would be expected to extinguish over training according to reinforcement learning theory. To explain its persistence, we propose that impulsive feeder approach is a behavior wherein an animal approaches areas where reward is

delivered, but does not take into account task-related responses necessary to be rewarded. This may be related to goal tracking in Pavlovian approach, which also increases with acute AMPH (Holden and Peoples, 2010). Goal tracking is a conditioned stimulus-response action wherein an animal approaches areas or objects associated with goals, which in our task is represented by the sucrose water dispensing feeders. However, this reflexive behavior ignores the behavior sequence that is required in the task and will therefore not be rewarded. The role of VS in Pavlovian approach has been well established, with lesions of VS core impairing performance on conditioned stimulus-based autoshaping tasks both during and after training (Cardinal et al., 2002). Furthermore, experimental lesions of dopaminergic input to VS causes severe impairments in Pavlovian approach tasks (Parkinson et al., 2002), suggesting that dopamine is necessary for the expression of this behavior. Therefore, increased monoaminergic tone in the VS by on-board AMPH may increase the expression of this behavior. Moreover, the increased spine density following sensitization may thus lead to an increased prevalence of VS-mediated control of behavior, which would be expressed as increased impulsive feeder approach in our task.

Some of the behavioral and neurochemical changes resulting from AMPH sensitization are similar to those associated with schizophrenia, leading some to propose sensitization as a model of schizophrenia (Featherstone et al., 2007). If this is a valid model, we would predict a shift of control away from DMS (caudate in primates) and toward VS (nucleus accumbens) and DLS (putamen) to be associated with the pathology. We speculate that the disruption of neural signaling in the prefrontal cortex associated with schizophrenia in humans (Rolls et al., 2008) and in animal models of the illness

(Gruber et al., 2010; Niwa et al., 2010; Molina et al., 2014) could be a driving force for a similar shift. This idea is weakly supported by a recent study using resting-state functional magnetic resonance imaging, which demonstrated a correlation between increased activity in dorsal or ventral striatum and increased positive or negative symptoms of schizophrenia (Sorg et al., 2013). We predict that such activity shifts would lead to increased lose-shift behavior and/or Pavlovian approach analogues in patients with schizophrenia.

Although lose-shift is a relatively simple behavior, it likely has consequences for behavioral flexibility in more complex tasks. For instance, increased lose-shift responding may promote apparent flexibility in tasks solvable by avoidance of particular actions after reinforcement contingency depreciation. The basis of lose-switch responding in the DLS, however, suggests that it is limited in solving complex or goal-directed problems as compared to other neural circuits (Daw et al., 2005; Balleine and O'Doherty, 2010; Gruber and McDonald, 2012). Lose-shift as well as impulsive feeder approach are both sub-optimal behaviors on the task that reduce the total amount of reward the animal can collect during a session; the former because it is predictable and can be exploited by the computer opponent, and the latter because the animal spends time engaging in an activity that is never reinforced. We suggest that the prevalence of these suboptimal behaviors indicates they are intrinsic, engaged reflexively in some circumstances, depend on dissociated striatal regions, and may be adaptive in some ethological contexts such as foraging.

2.5 Conclusion

We have demonstrated that acute injections of AMPH reduce lose-shift responding, whereas sensitization results in an increased lose-shift responding in the same task. The former could reflect attenuation of worse-than-expected signals by dopamine, whereas the latter may reflect a shift of behavioral control to DLS, as proposed to occur during drug addiction. We further discovered that the approach to multiple feeders is increased by acute AMPH delivered systemically or infused into VS, but not in DLS. We argue that these simple behavioral responses are reflexively expressed by dissociated striatal circuits, and both will influence cognitive flexibility in more complex tasks.

3. Acute Δ -9-tetrahydrocannabinol reduces loss sensitivity in a binary choice task while preserving global behavioral flexibility

Abstract:

The main psychoactive ingredient in marijuana is THC, which has potent effects on decision-making. Here we demonstrate that acute THC administration reduces sensitivity to reward omission in rats, as revealed by a decrease in lose-shift responding during a competitive binary choice task and in a multiple reversal learning paradigm. Other aspects of responding, such as motor output and measures of motivation, were unaffected. Because THC has been proposed to impair cognitive flexibility, we also analyzed choice perseveration and stereotyped patterns of responding in these tasks by quantifying the randomness of choice sequences. THC administration did not affect choice randomness of rats performing the competitive binary choice task, wherein random responding is optimal. In contrast, THC did reduce choice randomness and increase choice perseveration compared to controls in the multiple reversal learning paradigm; however these adaptations are beneficial in this task and should not be considered impairments. Furthermore, THC did not impair the ability of animals to shift response biases to track reversals of reward probabilities in the task. This null effect of THC on cognitive flexibility in the current task contrasts other reports; we speculate that the lack of sensory cues in the present task may contribute to this discordance. Together, these data suggest that THC will more strongly impair tasks in which rats utilize lose-shift strategies rather than other reinforcement-driven response adaptations.

3.1 Introduction

Marijuana and its main psychoactive ingredient THC, are one of the most commonly used illicit drugs in the world, with an estimated 13.1 million people exhibiting cannabis dependence as of 2010 (Degenhardt et al., 2013). With usage increasing in many countries (Hall and Degenhardt, 2007; Hasin et al., 2015) and various governments considering legalizing its use, there is a pressing need to more fully understand the effects of THC on decision-making and cognition. Previous studies indicate deficits in executive function, verbal and visual memory, and visuoperception following cannabis use in humans (Bolla et al., 2002). Rodent models have also indicated impaired working memory (Varvel et al., 2001; Fadda et al., 2004) and spatial memory (Lichtman and Martin, 1996) following THC administration. Furthermore, THC causes impairments in reversal learning in macaques (Wright et al., 2013), as well as reversal learning and intradimensional set shifting impairments in rats (Egerton et al., 2005; Sokolic et al., 2011). These findings support the proposal that THC impairs cognitive flexibility. Two brain regions involved in regulating cognitive flexibility are orbitofrontal cortex (OFC), and lateral prefrontal cortex (LPFC). Lesions of primate OFC impair reversal learning (Dias et al., 1996), while LPFC lesions selectively impair extradimensional set shifting (Dias et al., 1996). Comparable results are also found with rat OFC (Schoenbaum et al., 2002) and medial prefrontal cortex lesions (Birrell and Brown, 2000). The tasks used in these studies, as is common among those used to measure choice flexibility, likely recruit several learning and memory systems in order to maintain dynamic associations among specific cues, actions, and reinforcements so as to engage appropriate responses.

A common methodological feature among these reversal learning and set shifting tasks is that they utilize sudden and unexpected changes in reward contingencies to assess how animals adapt behaviour from trial-and-error experience. A physiological mechanism thought to be important for such reinforcement-driven response adaptation is the RPE signaling properties of midbrain dopamine neurons, wherein dopamine neuron firing briefly increases following unexpectedly good reinforcements (rewards), but decrease firing following unexpectedly poor reinforcements such as reward omission (Schultz et al., 1997; Fiorillo et al., 2003; Roesch et al., 2007). Drugs or diseases that affect dopamine transmission are therefore expected to impair normal reinforcement learning, which has been demonstrated in many species and tasks (Jentsch et al., 2002; Frank et al., 2004; Gradin et al., 2011; Wong et al., 2016).

Similar to other drugs of abuse, THC increases midbrain dopamine neuron firing (French et al., 1997) and increases dopamine concentration in VS shell (Tanda et al., 1997). THC acts as an agonist at Cannabinoid Receptor 1 (CB1), which is G_i coupled and acts as an inhibitory presynaptic regulator of neurotransmitter release (Svíženská et al., 2008). CB1 is located on inhibitory afferents to dopamine neurons (Sañudo-Peña et al., 1999), and so it is likely that that CB1 agonism reduces inhibition of dopamine neurons to increase their firing rate and transmitter release. The increased levels of dopamine evoked by THC via CB1 agonism are therefore expected to suppress the pause in dopamine neuron firing purportedly signaling a negative RPE following a loss (Schultz et al., 1997). An analogous process likely occurs with other dopamine-acting drugs. We recently showed that AMPH selectively attenuates the propensity of rats to switch choices after a reward omission, a phenomenon termed lose-shift responding (Wong et al., 2016). The

attenuation of this behavior has important implications in tasks that rely on reward omission to motivate changes in behavior. We hypothesized that the reported impaired cognitive flexibility by acute THC may also involve reduced lose-shift responding due to supraphysiological dopamine levels.

To test this hypothesis, we systemically injected rats with THC and analyzed specific features of reinforcement-driven response adaptation during two choice tasks. In particular, we sought to differentiate between effects on lose-shift responding and the more general cognitive flexibility needed to track reversals of asymmetric reward probabilities among binary response options. We found that systemic THC administration attenuated the propensity of rats to shift responses after reward omission, but did not impair their ability to flexibly reverse choice preference to track uncued changes in reward probability. These data suggest that acute CB1 agonism specifically reduces lose-shift responding while not affecting the flexibility needed for reversing choice preference based on reinforcements.

3.2 Methods

3.2.1 Subjects

Subjects were 21 adult Long-Evans rats (bred in-house) weighing 200-250 g. Animals were pair-housed in a climate-controlled vivarium under a 12:12 hour light:dark cycle (lights on 7:30 a.m.). Animals had restricted access to water (one hour) on behavioral testing days, but otherwise had ad libitum access to food and water. All procedures were approved by the University of Lethbridge Animal Welfare Committee, following the guidelines of the Canadian Council on Animal Care.

3.2.2 Behavior apparatus

Behavioral testing was performed in aluminum operant chambers as described previously (see Skelin et al. (2014)). Briefly, animals were placed in the operant chamber for one-hour sessions. Trials were self-paced, and initiated by the animal performing a nose-poke into the central port. Following 150ms of nose-poke entry, a tone (6 KHz) was presented to indicate satisfactory nose-poke, and the animal could then select one of the two adjacent sucrose delivery feeders. If the correct feeder was chosen, a reward (60 μ L of 10% sucrose solution) was delivered after a 500ms delay. If the incorrect feeder was chosen, no sucrose was delivered, the house-light illuminated, and the two panel lights extinguished. The state of the lights then reverted (house->off; panel->on) after 200-400ms. This change in lighting served to indicate that reward was not forthcoming, and was of sufficiently short durations such that it terminated by the time the animals could return to the central poke port. This means there was no 'time-out' associated with reward omission. Once a feeder was chosen, or if no feeder was chosen in the 15s following a nose-poke, the trial ended and the animal had to return to the central port to initiate a new trial.

3.2.3 Task 1: Matching Pennies (MP) Task

All trials were rewarded and no barriers were present in the first training session to facilitate task acquisition. In the second training session, animals were rewarded on 50% of the trials regardless of feeder choice. In all subsequent sessions, reinforcement was controlled by an algorithm that attempted to minimize the number of rewards given to the animal by predicting which feeder it would select. This was done by examining the

choices and reinforcements from the previous four trials (Lee et al., 2004). If either feeder was selected at a greater than chance rate, it would be unrewarded for the upcoming trial. In doing so, the competitive mode implements the classic ‘Matching Pennies’ task. Parallel barriers positioned between the central port and feeder wells were added to introduce a choice cost and discourage feeder bias due to body position. Increasingly longer barriers (4.0, 8.5, 13.5 cm) were introduced during consecutive days of training. Rats were trained until they completed two consecutive sessions of at least 150 trials with the long barriers. All subsequent training and testing sessions were run with the long barriers.

3.2.4 Task 2: Probability Reversal (PR) Task

A second behavioral task was performed in the same operant chambers. Trial initiation was self-paced and began with a nose poke in the central port. A non-informative tone then prompted the animal to select a sucrose delivery feeder. Correct feeder choices were rewarded with sucrose solution after a 500ms delay. No reward was given for incorrect choices. In contrast to the MP task, the probability of receiving a reward at a particular feeder was fixed over blocks of 60 trials to either a high or low reward probability. Each consecutive block of trials in a session would reverse these feeder reward probabilities. For example, the left feeder would be the high reward probability side on trials 1-60, and would then reverse to become the low reward probability side for trials 61-121. Animals were trained on this task for six sessions, with the high/low reward probability increasing every two sessions from 0.6/0.4 to 0.7/0.3, and finally to 0.8/0.2.

3.2.5 Drug preparation and injections

For Experiment 1, THC (Cayman Chemicals, Ann Arbor: Michigan) was dissolved into a 1:1:1:16 solution of THC:ethanol:Cremaphor EL:sterile saline (0.9%). THC solution was delivered by IP injection at one of three dosages (0.5mg/kg, 1.0mg/kg, or 2.0mg/kg). For Experiment 2, animals were given IP injections of THC (2.0mg/kg) or vehicle. Injection sites were rotated and alternated sides to minimize irritation.

3.2.6 Experiment 1: Acute effects of THC on the MP task

After initial shaping (9 daily sessions), animals (n=11) were randomly divided into four groups to receive acute THC in a counterbalanced block design. Injections were administered 30 minutes prior to testing on the behavioral task over a period of 8 days using the following schedule: vehicle, injection 1, no injection, injection 2, no injection, injection 3, no injection, and injection 4. The initial vehicle injection was to habituate animals to the procedure and was not included in the analysis. Injection days consisted of one of the three THC dosages or vehicle.

3.2.7 Experiment 2: Acute effects of THC on PR task

After initial training, animals (n=10) were randomly divided into two groups (n=5, n=6) and received either THC or vehicle 30 minutes before testing in a counterbalanced design with reward probabilities for testing sessions set at 0.8/0.2. The injection schedule over days was as follows: vehicle, no injection, injection 1, no injection, injection 2. The initial injection of vehicle was to habituate animals to the process and was not included in the analysis.

3.2.8 Behavior analysis

Data were analyzed using Matlab version 2013a (Mathworks, Massachusetts) and Graphpad Prism version 6 (Graphpad, California), while SPSS version 21 (IBM, New York) was used to test sphericity. To minimize potential confounds, all trials in which the animal sampled both feeders on the previous trial (IFA) were omitted from the analysis of lose-shift and win-stay probabilities. For Experiment 1 (Acute effects of THC on the MP task), we examined several features of responding including: lose-shift, win-stay, reaction time, number of trials, response time, response entropy, intertrial interval, and impulsive feeder sampling. Statistical significance was determined by one-way repeated measures ANOVAs. The Greenhouse-Geisser correction has been applied to these calculations when the sphericity assumption was violated, as determined by Mauchly's test of Sphericity ($\alpha=0.05$). Lose-shift responding is a measure of loss sensitivity wherein animals chose to shift feeder choice from that of the previous trial if they failed to receive reinforcement on that trial. Win-stay is the re-selection of the feeder on the current trial that provided a reward on the previous trial. Response entropy is a measure of the randomness of an animal's choice sequence (Skelin et al., 2014), with higher values representing more unpredictable responses. This is determined by calculating the Shannon entropy (see 3.4 Discussion for more details) of the animal's previous four trials (Miller, 1955). One animal was removed from Experiment 1 due to being an extreme outlier on most behavioural measures, as determined by Grubb's test ($\alpha=0.05$). In Experiment 2 (Acute effects of THC on the PR Task), the same measures were analyzed using paired t tests because only one level of the drug was used. To quantify changes of behavior within each block of 60 trials, we collapsed across blocks and binned trials in

blocks of ten. This created six distinct epochs within the blocks. Two-way repeated measures ANOVAs were used to determine statistical significance of drug and bin number (experience) within blocks.

3.3 Results

3.3.1 Experiment 1 (Acute effects of THC on the MP task)

We trained 10 female Long-Evans rats to perform a competitive binary choice task, sometimes referred to as ‘Matching Pennies’ (see Methods). We first assessed the potential effects of THC on motivation and motor output. Drug had no effect on the number of trials performed in 1 hour sessions ($F(3, 30)=1.21$, $p=0.32$; Fig 2.A). Although not statistically significant, reaction time ($F(3,30)=2.78$, $p=0.11$; Fig 2.B) and response time ($F(3,30)=2.628$, $p=0.07$) tended to increase as drug dose increased. THC had no effect on the intertrial interval ($F(3, 30)=1.65$, $p=0.22$; Fig 2.C), or the number of anticipatory licks at the feeder ($F(3, 30)=0.323$, $p=0.80$). Together, these results suggest that THC had a minimal effect on motor output and motivation. Although random responding with insensitivity to loss and reward on the previous trial is the optimal strategy in this task, we previously discovered that rats persistently used a lose-shift and win-stay strategy (Skelin et al., 2014). We found here that systemic THC significantly decreased lose-shift responding in a dose-dependent manner ($F(3, 30)=5.34$, $p=0.005$; Fig 2.D), similar to the effect of systemic AMPH (Wong et al., 2016). Win-stay responding also significantly decreased with increasing THC dose ($F(3,30)=3.52$, $p=0.03$; Fig 2.E), while the tendency of animals to impulsively approach feeders without first committing a nose-poke was unaffected ($F(3,30)=1.57$, $p=0.22$). THC did not affect the randomness of choice as measured by the response entropy ($F(3,30)=0.598$, $p=0.57$; Fig 2.F), suggesting

that it is the sensitivity of choice on reinforcement that is affected by the drug rather than an overall change in response pattern. Furthermore, THC had no effect on the percentage of rewarded trials ($F(3,30)=1.34, p=0.28$), supporting the idea that drug caused no impairment on decision-making.

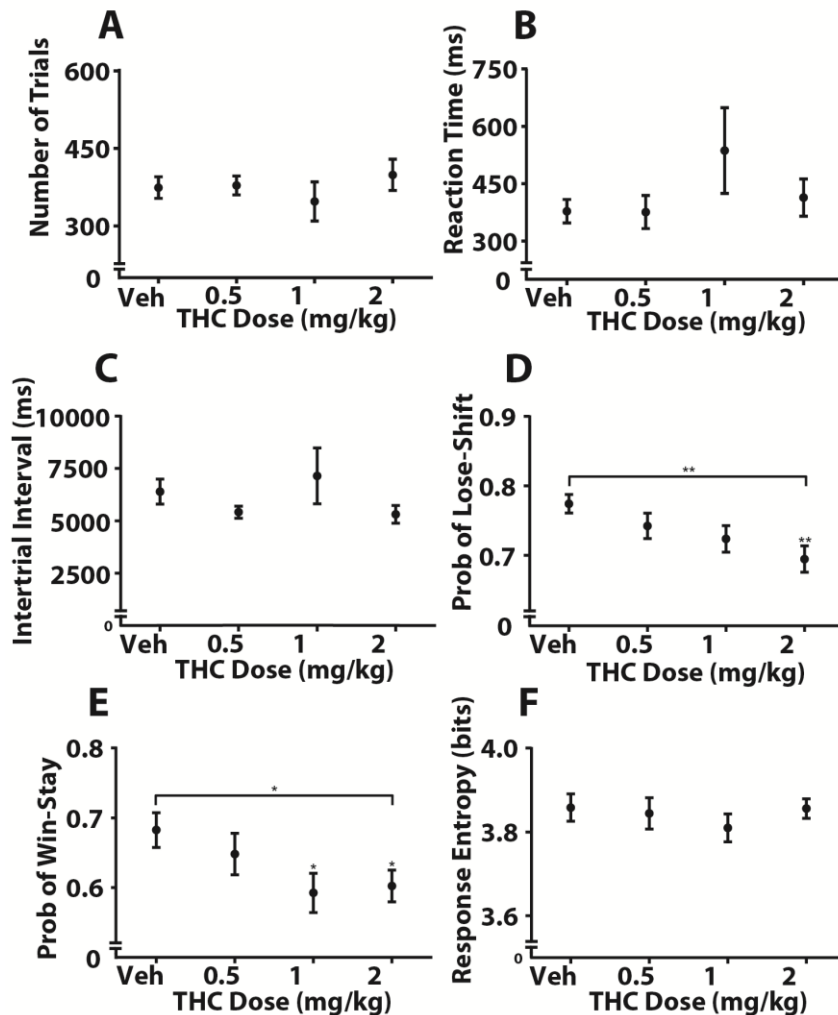


Figure 7. Effects of acute THC administration on behavioral measures in the MP task. (A) Drug had no effect on the number of trials performed. (B) Reaction time was unaffected by THC. (C) Intertrial interval time was also unaffected by THC. (D) Probability of lose-shift responding decreased in a dose-dependent manner. (E) Probability of win-stay responding, which was also reduced in a dose-dependent manner. (F) THC had no effect on response entropy. Statistical significance for repeated measures-ANOVAs are indicated by ‘*’ for $p < 0.05$, and ‘**’ for $p < 0.01$. Asterisks above individual columns indicates group means that were significantly different from the control (Vehicle) mean according to Dunnett's post hoc test with $\alpha = 0.05$.

3.3.2 Experiment 2 (Acute effects of THC on the PR Task)

We next sought to determine if THC affected animal's cognitive flexibility by assessing their ability to alter choice strategy in response to changing reward contingencies. We trained 10 male Long-Evans rats to perform the PR task (see Methods) and compared behavior on and off THC at the most extreme probabilities (0.8/0.2). To maximize reward in this task, rats must quickly learn which feeder has a higher probability of reward and show the flexibility to reverse this bias when a new block of trials begins and feeder reward probabilities reverse. Unlike the first task, in which the optimal response strategy is random choice, the rats should develop a strong bias for choosing the feeder resulting in a higher probability of reward. Consequently, they could earn reward on nearly 80% of the trials by choosing optimally, but would earn reward on 50% of trials by choosing randomly. In practice, rats acquired reward on $66 \pm 3\%$ of trials (see Fig 2), indicating that they did adapt to exploit asymmetries in the reward probabilities.

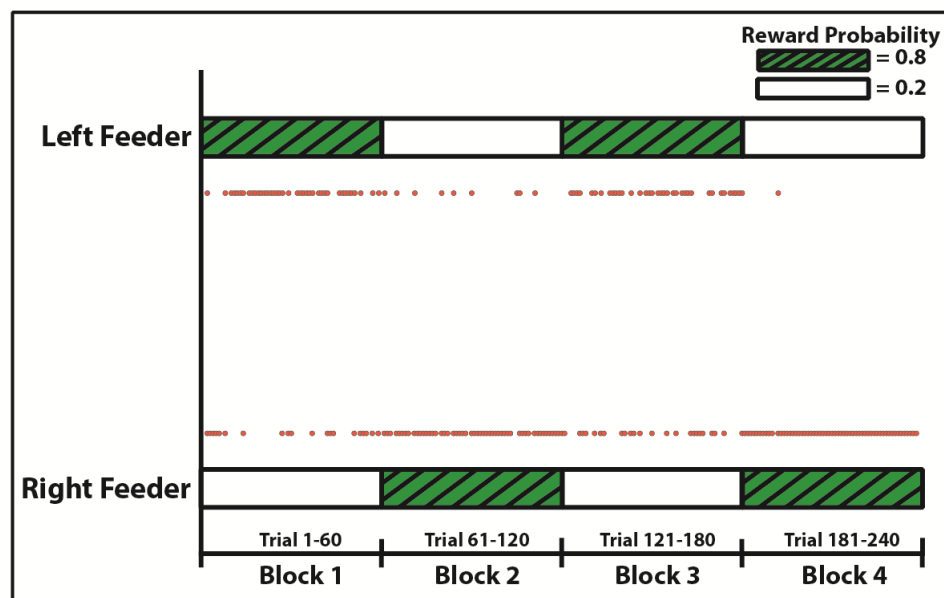


Figure 8. Representative data (240 trials over 4 blocks) of a rat's feeder choices during the PR task. Each orange dot represents an individual trial and demonstrates shifting feeder preference in response to changing reward contingencies.

Similar to Experiment 1, drug administration did not significantly affect the number of trials per session ($t(10)=1.65$, $p=0.13$; Fig 4.A), intertrial interval ($t(10)=1.23$, $p=0.25$; Fig 4.B), or pre-reward licks ($t(10)=0.07$, $p=0.95$), further suggesting no gross impairment in motor control or motivation. However, reaction time ($t(10)=2.50$, $p=0.03$) and response time ($t(10)=2.43$, $p=0.04$) significantly increased with THC administration, which is consistent with the reported effects of high doses of the drug in previous studies (Sañudo-Peña et al., 1999). Similar to our results from Experiment 1, systemic THC significantly decreased lose-shift responding in a dose-dependent manner ($t(10)=3.47$, $p=0.007$; Fig 4.C). In contrast to Experiment 1, win-stay responding significantly *increased* following THC injection with increasing drug dose ($t(10)=2.37$, $p=0.04$; Fig 4.D), while impulsive feeder approach was again unaffected by drug ($t(10)=0.059$, $p=0.95$). THC administration significantly reduced response entropy ($t(10)=3.83$, $p=0.004$; Fig 4.E). This can likely be explained by the fact that THC significantly increased the probability of the animal returning to the previous feeder choice regardless of choice outcome ($t(10)=3.38$, $p=0.008$; Fig 4.F).

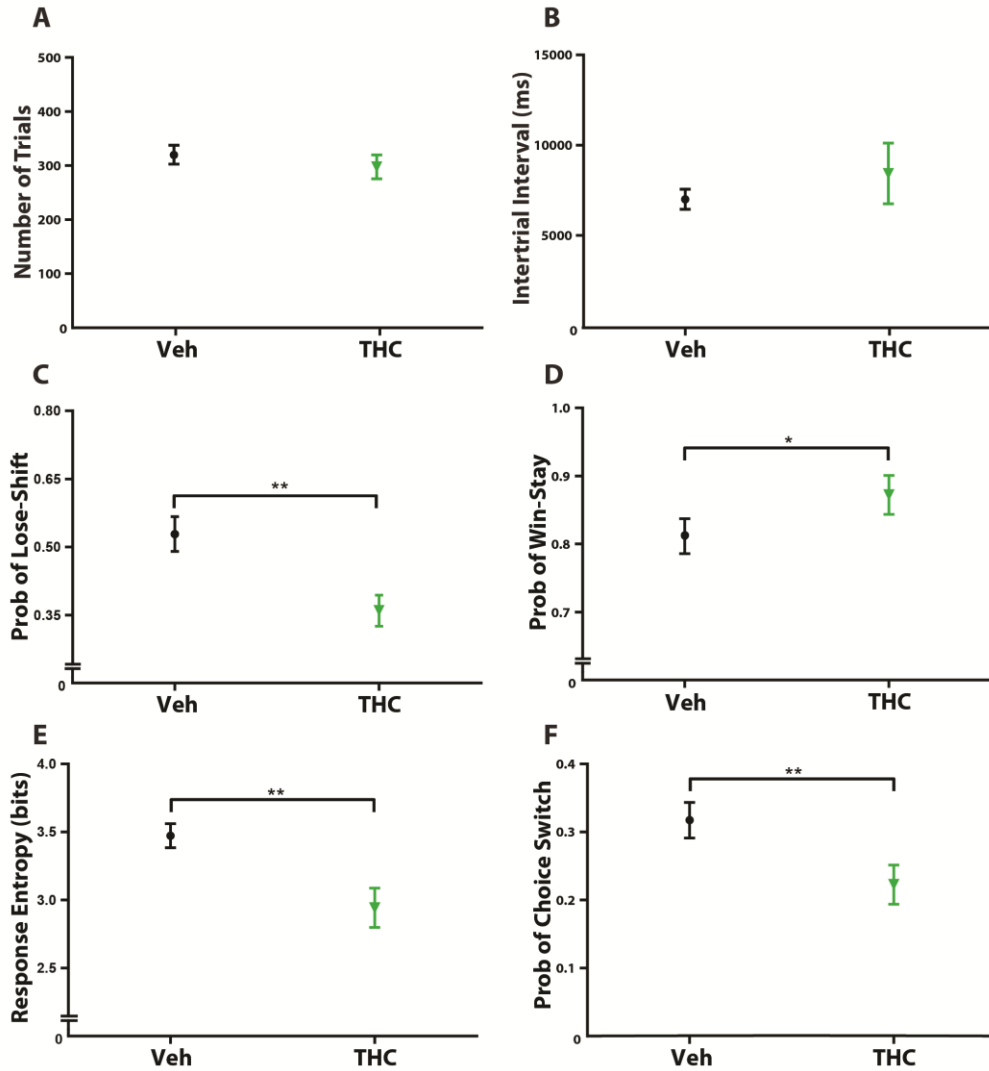


Figure 9. Session-averaged effects of acute THC in the PR task. (A) THC did not affect the number of trials performed. (B) No effect of THC was found on the time in between trials. (C) Probability of lose-shift responding, which was decreased by systemic THC. (D) Probability of win-stay responding, which was decreased by THC. (E) THC significantly reduced response entropy. (F) The probability of a switch from the previous choice independent of reinforcement was significantly reduced by THC. Statistical significance is indicated by ‘*’ for $p < 0.05$, and ‘**’ for $p < 0.01$ for paired t-tests.

We next wanted to determine whether our measures of choice varied across trials within a block. Optimal performance on the task requires animals to reverse choice bias when reward probabilities reverse at the start of each block. To investigate this and other behavioral measures, we collapsed the data across blocks, and aggregated the 60 trials in

each block into six equal bins. Our data suggest that animals adapted to the changing reward probabilities; as they were more likely to receive a reward in the later trial bins within each block (main effect of bin number: $F(5,45)=21.5$, $p<0.0001$; Fig 5.A) while drug had no effect on the probability of receiving a reward (main effect of drug: $F(1,9)=0.003$, $p=0.95$; Fig 5.A). Animals were also less likely to switch their feeder choice from the previous trial regardless of outcome as blocks progressed (main effect of bin number: $F(5,45)=3.16$, $p=0.02$; 5.B), and this was further decreased with drug (main effect of drug: $F(1,9)=12.6$, $p=0.006$; 5.B). As with our session-average analysis, drug had a significant effect on lose-shift responding ($F(1,9)=17.6$, $p=0.002$; Fig 5.C), with drug reducing the probability of this behavior while bin number had no effect ($F(5,45)=2.31$, $p=0.06$; Fig 5.C). The drug also had a weakly significant effect on win-stay ($F(1,9)=5.28$, $p<0.05$; Fig 5.D), while bin number had a strong effect on win-stay ($F(5,45)=3.82$, $p=0.006$; Fig 5.D), with both THC and Vehicle groups increasing this response as blocks progressed.

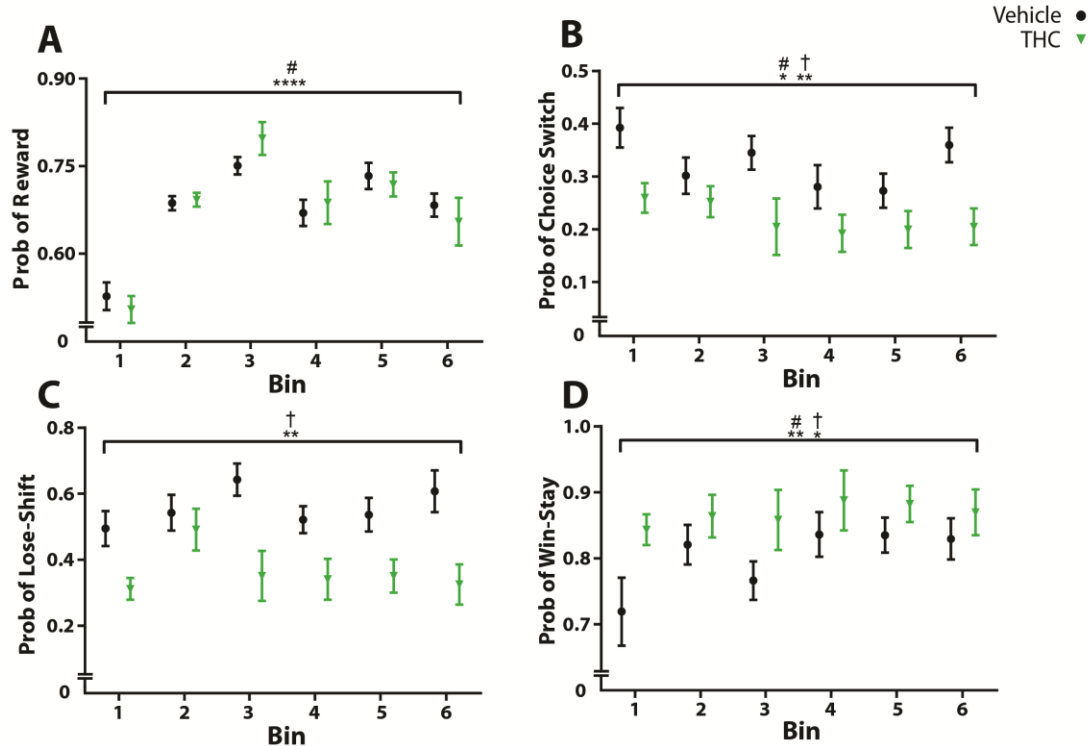


Figure 10. Within-block effects of THC on behavioral measures in the PR task. (A) THC had a significant main effect on the occurrence of lose-shift. (B) Drug and bin number both had significant main effects on win-stay responding. (C) The probability of reward on any particular trial was strongly affected by bin number. (D) The probability of a switch from the previous choice independent of reinforcement was significantly affected by both drug and bin number. Significance as determined by a two-way RM-ANOVA is indicated by ‘*’ for $p < 0.05$, ‘**’ for $p < 0.01$, and ‘****’ for $p < 0.0001$. A main effect of drug is indicated by ‘†’, while a significant main effect of bin number is indicated by ‘#’.

3.4 Discussion

The purpose of this study was to investigate the effects of on-board THC in two simple reward-based decision-making tasks. The first task (Matching Pennies) is an uncued competitive binary choice task wherein random responding is the optimal solution against a rational opponent. The second task (Probability Reversal) has reward probabilities at each feeder fixed at a high or low value within a block of trials, and these probabilities reverse after the uncued beginning of each block. This task requires that

animals rapidly alter their choice strategy each block to bias responding towards the high probability feeder. By including both these tasks, we were able to determine if THC affects specific decision strategies, such as lose-shift, win-stay or random responding, when such strategies are adaptive or maladaptive. This provides new insight into the effects of this drug on cognitive flexibility.

Table 2. Comparison of behavioral measures influenced by THC in the Matching Pennies and Probability Reversal tasks. Strongly significant effects are denoted with two arrows, while significant effects are denoted with one arrow. Arrow directions indicate directionality of differences as compared to Vehicle injections. Non-significant effects are denoted with ‘-’.

Behavioral Measure	Matching Pennies	Probability Reversal
Lose-Shift	↓↓	↓↓
Win-Stay	↓(weak)	↑(weak)
Impulsive Feeder Approach	-	-
Response Entropy	-	↓↓
Number of Trials	-	-
InterTrial Interval	-	-
Reaction Time	-	↑
Response Time	-	↑

THC administration weakly increased reaction time and response time compared to vehicle in the PR Task, and showed a non-significant trend to increase these in the MP task (see Table 2). The multiple doses of drug administered in the MP task exhibited an inverted U dose-response curve with increases in both measures with the low and medium dose, and decreases with the highest dose (data not shown). This is consistent with other studies showing a triphasic motor effect from systemic THC administration (Sañudo-Peña et al., 2000). This manifests as decreasing motor activity with low doses of THC, while higher doses stimulated motor activity up to the point of cataplexy with the highest doses.

We also found that lose-shift responding decreased in a dose-dependent manner in both tasks (see Table 2). This is consistent with our previous studies using AMPH (Wong et al., 2016), and provides further evidence supporting the idea that drugs can impair learning mechanisms dependent on dopaminergic signaling. Interestingly, lose-shift responding was stable across trials in the PR task with and without drug on-board. This leads us to hypothesize it is a reflexive, innate behavior rather than learned. Other work from our lab has shown that DLS is the likely locus of lose-shift behavior (Skelin et al., 2014). Dorsolateral striatum is innervated by dopamine neurons from substantia nigra (Haber et al., 2000; Joel and Weiner, 2000) and is historically considered a slow learning system associated with habit formation (Yin et al., 2004). However, recent optogenetic work has shown a similar functional role of dopamine neurons in substantia nigra and ventral tegmental area projections to the DMS and VS (Ilango et al., 2014).

A possible link connecting DLS with the cannabinoid system is the pattern of CB1 mRNA expression in striatum. CB1 mRNA expression has a ventromedial to dorsolateral gradient in striatum, with greater expression towards the dorsal regions of lateral striatum (Herkenham et al., 1991; Julian et al., 2003; Martín et al., 2008). This suggests a potential involvement of CB1 receptors in DLS in affecting lose-shift responding. It should be noted that while CB1 mRNA is expressed primarily in DLS, immunohistochemistry experiments have demonstrated that the CB1 receptor itself is primarily located on the corresponding MSN afferents located in the substantia nigra (Sañudo-Peña et al., 1999). This leaves DLS CB1 receptors ideally placed to presynaptically modulate MSN-driven inhibition of substantia nigra dopamine release and influence dopaminergic prediction errors. Immediate early gene expression also supports the role of DLS in altering choice

in a reward task (Egerton et al., 2005). The authors found that *c-fos* expression in DLS was correlated with performance on the first reversal of a reversal learning task, while *ngfi-b* expression was strongly correlated with performance on the third reversal. These results suggest that neural activity in DLS is linked to altering choice on a short timescale.

The idea of fluctuating dopamine levels in areas such as DLS rapidly influencing choice is consistent with the RPE hypothesis of dopamine. Because both THC and AMPH increase dopamine levels, it should not be surprising that they have a similar effect to reduce lose-shift behavior. However, these drugs have distinct mechanisms of action, so discrepancies in their effects on behaviour make sense. We observed opposing effects of AMPH and THC on several behavior measures in our MP task. Our previous work found that AMPH has no effect on win-stay responding (Wong et al., 2016) while THC increases this behavior in the PR task and decreases it in the MP task. In contrast, the tendency to approach feeder to seek out of trial rewards (IFA) was unaffected by THC, whereas both systemic AMPH and infusions into VS increase the prevalence of this behavior.

Similar to AMPH, THC did not affect response entropy (see Table 2) in the MP task. Entropy assesses the randomness of choice sequence (in this study length = 4) over a behavior session (Miller, 1955). Entropy is maximized when animals choose a feeder randomly, with no pattern or bias. Conversely, entropy is reduced when they show a bias towards a feeder or particular pattern of responding (e.g. L, R, L, R). In the MP task, random responding is the optimal solution. While THC did reduce lose-shift responding, it had no effect on overall entropy in this task. This suggests that rats retained the ability to flexibly choose either well with no apparent patterned behavior.

In contrast to the MP task, THC did reduce entropy on the PR task (see Table 2), due to the fact that THC-treated animals were less likely to switch feeder choice after losses (decreased lose-shift) or wins (increased win-stay). Rats should repeatedly choose the higher probability option regardless of wins or losses in this task, so these effects from THC do not impair performance. This is supported by the fact that drug had no effect on the amount of reward obtained in the task. This furthermore indicates that THC did not impair the rats' ability to flexibly alter choice bias to track reward reversals.

The present data are contrary to conclusions presented in previous studies; we must therefore consider possible confounds to our experiment. One possibility is that sex differences influence our results, although we argue this is unlikely. We have controlled for sex in Long Evans rats performing the MP task and found only minor differences unrelated to the main findings here (manuscript in preparation by colleague). Moreover, we expect little to no effect of estrous cycle because an extensive meta-analysis of behavioral and physiological traits found equal variability in both sexes (Prendergast et al., 2014), even when including estrous cycle phases. Our tasks also vary from others (Egerton et al., 2005; Wright et al., 2013) in that rats are not given informative cues during the task about reinforcement outcomes. It is possible that without such cues, the orbitofrontal and prefrontal cortices have less influence on task performance and so their drug-induced dysfunction will not manifest in behaviour. This would mean rats are able to remain flexible in their decision-making. Furthermore, rats do sometimes receive reward on the low-probability feeder. This differs from most reversal learning tasks, which tend to provide no reward (or aversive reinforcement such as quinine) on the previously optimal choice after a reversal has occurred.

As with all systemic pharmacology experiments, we also cannot ascertain which brain structures are affected. While CB1 expression is greatest in regions such as the striatonigral pathway, amygdala, and hippocampus, moderate expression also occurs in the cortex (Svíženská et al., 2008). Although we have not found an effect of electrolytic OFC lesions on lose-shift responding or response entropy (unpublished data), we cannot rule out a potential effect of drug on regions involved in reversal learning (OFC) and set shifting (LPFC). Nonetheless, our results are consistent with studies in humans examining the effect of acute THC on performance in the Wisconsin Card Sorting Task (Weinstein et al., 2008). The researchers found that THC-induced increases in errors were not due to impairments in flexibly tracking category changes (shape, color, or number), but rather caused by increases in what they termed ‘non-perseverance’ errors (errors within categories) indicating inattention. As the Wisconsin Card Sorting Task is also an uncued decision-making task, it supports the idea that THC does not reduce cognitive flexibility by promoting inflexible patterns of responding. Other human studies have also indicated no impairment of cognitive flexibility (Hart et al., 2001), although the subjects in this study were chronic drug users and may have learned to compensate for their impairments. Together, these data reveal that THC has complex and subtle effects on decision-making that remain incompletely understood.

3.5 Conclusion

Our data indicate that THC reduces the effect of reward omission on trial-by-trial basis, possibly by attenuating the normal pause in dopamine transmission that follows losses. We have shown that this occurs both in competitive choice tasks where

insensitivity to losses and wins is the optimal strategy, and in a reversal learning paradigm where avoiding losses and chasing wins is the optimal strategy. We also provide evidence that THC does not appear to impair the cognitive flexibility needed to rapidly adapt to changing reward contingencies, or cause an increase in non-optimal, patterned behavior in these tasks. These data suggest that, at least in some tasks, THC does not impair global decision-making strategies, but does selectively reduce avoidance of choices associated with recent losses.

4. General Discussion

In these experiments we sought to investigate how specific elements of response strategies in reward-based decision-making could be influenced by drugs of abuse. We found that both AMPH and THC reduced lose-shift responding in rats, suggesting reduced loss sensitivity. This may be caused by an attenuation of the negative RPE associated with loss that would normally promote a change in decision-making. This effect was robust, occurring both in the MP task with acute administration of either AMPH or THC, and with THC in the reversal learning paradigm, suggesting it is not a task-specific or drug-specific effect.

Our examination of response entropy suggests this behavior is distinct from global decision-making strategies. Our results indicate that entropy in the MP task was unaffected by systemic or intracerebral infusions of AMPH, or by systemic administration of THC. Conversely, systemic injections of THC did reduce the probability of switching feeder choice regardless of reinforcement and increased the prevalence of win-stay responding on the PR task. While these are forms of perseveration that would suggest impaired cognitive flexibility, they are optimal solutions on this task. This is highlighted by the fact that on-board THC did not reduce the amount of reward received during the task.

Because both THC and AMPH increase dopamine levels, it should not be surprising that they have a similar effect on behavioral measures. However, they act on different neurotransmitter systems and vary in their mechanisms of action and so would be expected that some of their effects differ. For example, AMPH also increases release of the other catecholamines (Sulzer et al., 2005) while THC will inhibit transmitter

release from any presynaptic terminals where CB1 is expressed. This has interesting implications for corticostriatal circuits involved in decision-making, which have an inverse relationship between CB1 expression in striatal sub-regions and the cortical regions innervating these sub-regions (Van Waes et al., 2012). Sub-regions with high levels of CB1 expression (DLS) have comparatively low expression on the cortical afferents to DLS. Conversely, regions with low levels of CB1 expression (DMS) have relatively high levels of CB1 expressed on cortical glutamatergic afferents into DMS. This leads us to speculate that THC may have differing effects on the amount of dopamine released into distinct striatal sub-regions. This provides a potential mechanism by which some behavioral measures in our tasks could show incongruent effects from THC and AMPH.

Behaviorally, we observed opposing effects of AMPH and THC on several measures in the MP task (see Table 2). Neither systemic or VS infusions of AMPH affected win-stay responding while THC increased this behavior in the PR task and decreases it in the MP task. In contrast, the tendency to approach feeders to seek rewards out of trials (impulsive feeder approach) was unaffected by THC, whereas both systemic AMPH and infusions into VS increase the prevalence of this behavior.

Table 3. Comparison of effects of systemic AMPH and THC on behavioral measures in the MP task. Very significant effects ($p < 0.001$) are denoted with three arrows, strongly significant effects ($p < 0.01$) are denoted with two arrows, and significant effects are denoted with one arrow ($p < 0.05$). Arrow directions indicate directionality of differences as compared to Vehicle injections. Non-significant effects are denoted with ‘-’.

	AMPH		THC
	Acute	Chronic	Acute
Lose-Shift	↓↓	↑	↓↓
Win-Stay	-	-	↓ (weak)
Impulsive Feeder Approach	↑↑↑	-	-
Response Entropy	-	-	-
Number of Trials	-	-	-
Intertrial Interval	↓↓↓	-	-
Reaction Time	↓↓	-	-
Response Time	-	-	-

The increases in IFA with infusions into VS but not DLS suggest this behavior is modulated by VS. Histology suggests that our AMPH infusions were into VS core, so it is possible that IFA is further localized to this area. It has been suggested that VS core promotes response vigor and approach behaviors (Gruber and McDonald, 2012), so it is possible this approach behavior is a form of goal tracking in Pavlovian approach. Goal tracking is a conditioned stimulus-response action wherein an animal reflexively approaches areas or objects associated with goals, which in our task is represented by the sucrose water dispensing feeders. However, this reflexive behavior ignores the behavior sequence that is required in the task and will not result in reward. Experimental lesions of VS core impair performance on conditioned stimulus-based autoshaping tasks both during and after training (Cardinal et al., 2002), suggesting on VS core has an ongoing role in Pavlovian approach. Furthermore, lesions of dopaminergic input to VS causes severe impairments in Pavlovian approach tasks, indicating that dopamine is necessary for the

expression of this behavior (Parkinson et al., 2002). Therefore, increased dopaminergic tone in the VS by infused AMPH may increase the expression of this behavior.

In contrast to acute administration of AMPH or THC, drug sensitization with AMPH increased lose-shift responding. We propose this may be due to increased behavioral control by DLS as a result of drug-induced plastic changes. Our findings of increased dendritic spine density in the region following chronic administration of AMPH concur with other findings in the literature (Robinson and Kolb, 1997; 1999; Jedynak et al., 2007). These results would also suggest that VS-driven behaviors such as impulsive feeder approach would also be increased following sensitization. While we did not find significantly increased impulsive feeder approach following sensitization, this measure approached significance and was greater in sensitized animals than controls on every day of testing. Taken together, we suggest these behaviors can serve as a barometer of behavioral control shifts in striatum that have been proposed to occur in the progression of addiction.

5. General Conclusion

Lose-shift responding has been observed in many tasks and species including humans (Wang et al., 2014), mice (Means and Fernandez, 1992), pigeons (Rayburn-Reeves et al., 2013), and honeybees (Komischke et al., 2002). The presence of this seemingly innate response in so many species suggests that it may serve an important evolutionary role, such as motivating animals to cease foraging in areas that do not have food and move on to new areas.

Lose-shift responding also has relevance to commonly utilized reward-based tasks used in research, such as reversal learning and set shifting. Apparent goal-directed changes in behavior in these tasks may be affected by the expression of this reflexive behavior, without actually altering cognitive flexibility. This has important implications in the interpretation of experiments utilizing reward-based tasks. Together, these experiments provide further evidence of the effects of acute and chronic administration of drugs of abuse on reward-based decision-making. This knowledge can assist in the design and interpretation of future experiments examining dopamine and decision-making, and will be of interest in creating effective drug rehabilitation programs.

6. References

- Alexander, G.E., DeLong, M.R., and Strick, P.L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annual Review of Neuroscience* 9, 357-381.
- Balleine, B.W., and O'Doherty, J.P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48-69.
- Bari, A., Theobald, D.E., Caprioli, D., Mar, A.C., Aidoo-Micah, A., Dalley, J.W., and Robbins, T.W. (2010). Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology* 35, 1290-1301.
- Birrell, J.M., and Brown, V.J. (2000). Medial frontal cortex mediates perceptual attentional set shifting in the rat. *The Journal of Neuroscience* 20, 4320-4324.
- Bolla, K.I., Brown, K., Eldreth, D., Tate, K., and Cadet, J. (2002). Dose-related neurocognitive effects of marijuana use. *Neurology* 59, 1337-1343.
- Cardinal, R.N., Parkinson, J.A., Lachenal, G., Halkerston, K.M., Rudarakanchana, N., Hall, J., Morrison, C.H., Howes, S.R., Robbins, T.W., and Everitt, B.J. (2002). Effects of selective excitotoxic lesions of the nucleus accumbens core, anterior cingulate cortex, and central nucleus of the amygdala on autoshaping performance in rats. *Behavioral Neuroscience* 116, 553.
- Castañé, A., Theobald, D.E.H., and Robbins, T.W. (2010). Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behavioural Brain Research* 210, 74-83.
- Caul, W.F., and Brindle, N.A. (2001). Schedule-dependent effects of haloperidol and amphetamine: multiple-schedule task shows within-subject effects. *Pharmacology Biochemistry and Behavior* 68, 53-63.

- Chang, Q., and Gold, P.E. (2004). Inactivation of dorsolateral striatum impairs acquisition of response learning in cue-deficient, but not cue-available, conditions. *Behavioral Neuroscience* 118, 383.
- Clarke, H.F., Robbins, T.W., and Roberts, A.C. (2008). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *The Journal of Neuroscience* 28, 10972-10982.
- Crombag, H.S., Gorny, G., Li, Y., Kolb, B., and Robinson, T.E. (2005). Opposite effects of amphetamine self-administration experience on dendritic spines in the medial and orbital prefrontal cortex. *Cerebral Cortex* 15, 341-348.
- Darvas, M., and Palmiter, R.D. (2009). Restriction of dopamine signaling to the dorsolateral striatum is sufficient for many cognitive behaviors. *Proceedings of the National Academy of Sciences* 106, 14664-14669.
- Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience* 8, 1704-1711.
- Degenhardt, L., Ferrari, A.J., Calabria, B., Hall, W.D., Norman, R.E., McGrath, J., Flaxman, A.D., Engell, R.E., Freedman, G.D., Whiteford, H.A., and Vos, T. (2013). The global epidemiology and contribution of cannabis use and dependence to the global burden of disease: results from the GBD 2010 study. *PLoS ONE* 8, e76635.
- Dias, R., Robbins, T., and Roberts, A. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature* 380, 69-72.
- Egerton, A., Brett, R.R., and Pratt, J.A. (2005). Acute Δ 9-tetrahydrocannabinol-induced deficits in reversal learning: Neural correlates of affective inflexibility. *Neuropsychopharmacology* 30, 1895-1905.
- Eiden, L.E., and Weihe, E. (2011). VMAT2: a dynamic regulator of brain monoaminergic neuronal function interacting with drugs of abuse. *Annals of the New York Academy of Sciences* 1216, 86-98.

- Everitt, B.J., and Robbins, T.W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience* 8, 1481-1489.
- Everitt, B.J., and Robbins, T.W. (2013). From the ventral to the dorsal striatum: Devolving views of their roles in drug addiction. *Neuroscience & Biobehavioral Reviews* 37, 1946-1954.
- Fadda, P., Robinson, L., Fratta, W., Pertwee, R.G., and Riedel, G. (2004). Differential effects of THC- or CBD-rich cannabis extracts on working memory in rats. *Neuropharmacology* 47, 1170-1179.
- Featherstone, R., Rizos, Z., Kapur, S., and Fletcher, P.J. (2008). A sensitizing regimen of amphetamine that disrupts attentional set-shifting does not disrupt working or long-term memory. *Behavioural Brain Research* 189, 170-179.
- Featherstone, R.E., Kapur, S., and Fletcher, P.J. (2007). The amphetamine-induced sensitized state as a model of schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 31, 1556-1571.
- Featherstone, R.E., and McDonald, R.J. (2004). Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a stimulus-response-based instrumental discrimination task, while sparing conditioned place preference learning. *Neuroscience* 124, 23-31.
- Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898-1902.
- Fleckenstein, A.E., Volz, T.J., Riddle, E.L., Gibb, J.W., and Hanson, G.R. (2007). New insights into the mechanism of action of amphetamines. *Annual Review of Pharmacology and Toxicology* 47, 681-698.
- Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., and Hutchison, K.E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences* 104, 16311-16316.
- Frank, M.J., Seeberger, L.C., and O'reilly, R.C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306, 1940-1943.

- French, E.D., Dillon, K., and Wu, X. (1997). Cannabinoids excite dopamine neurons in the ventral tegmentum and substantia nigra. *Neuroreport* 8, 649-652.
- Gradin, V.B., Kumar, P., Waiter, G., Ahearn, T., Stickle, C., Milders, M., Reid, I., Hall, J., and Steele, J.D. (2011). Expected value and prediction error abnormalities in depression and schizophrenia. *Brain*, 134, 1751-1764.
- Gruber, A.J., Calhoun, G.G., Shusterman, I., Schoenbaum, G., Roesch, M.R., and O'Donnell, P. (2010). More is less: a disinhibited prefrontal cortex impairs cognitive flexibility. *The Journal of Neuroscience* 30, 17102-17110.
- Gruber, A.J., and McDonald, R.J. (2012). Context, emotion, and the strategic pursuit of goals: Interactions among multiple brain systems controlling motivated behaviour. *Frontiers in Behavioral Neuroscience* 6:50.
- Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. *Journal of Chemical Neuroanatomy* 26, 317-330.
- Haber, S.N., Fudge, J.L., and McFarland, N.R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *The Journal of Neuroscience* 20, 2369-2382.
- Hall, W., and Degenhardt, L. (2007). Prevalence and correlates of cannabis use in developed and developing countries. *Current Opinion in Psychiatry* 20, 393-397.
- Hart, C.L., Van Gorp, W., Haney, M., Foltin, R.W., and Fischman, M.W. (2001). Effects of acute smoked marijuana on complex cognitive performance. *Neuropsychopharmacology* 25, 757-765.
- Hasin, D.S., Saha, T.D., Kerridge, B.T., Goldstein, R.B., Chou, S.P., Zhang, H., Jung, J., Pickering, R.P., Ruan, W.J., and Smith, S.M. (2015). Prevalence of marijuana use disorders in the United States between 2001-2002 and 2012-2013. *JAMA Psychiatry* 72, 1235-1242.
- Heien, M.L.A.V., Khan, A.S., Ariansen, J.L., Cheer, J.F., Phillips, P.E.M., Wassum, K.M., and Wightman, R.M. (2005). Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats. *Proceedings of the National Academy of Sciences* 102, 10023-10028.

- Herkenham, M., Lynn, A.B., Johnson, M.R., Melvin, L.S., de Costa, B.R., and Rice, K.C. (1991). Characterization and localization of cannabinoid receptors in rat brain: a quantitative in vitro autoradiographic study. *The Journal of Neuroscience* 11, 563-583.
- Holden, J.M., and Peoples, L.L. (2010). Effects of acute amphetamine exposure on two kinds of pavlovian approach behavior. *Behavioural Brain Research* 208, 270.
- Idris, N.F., Repeto, P., Neill, J.C., and Large, C.H. (2005). Investigation of the effects of lamotrigine and clozapine in improving reversal-learning impairments induced by acute phencyclidine and d-amphetamine in the rat. *Psychopharmacology* 179, 336-348.
- Ilango, A., Kesner, A.J., Keller, K.L., Stuber, G.D., Bonci, A., and Ikemoto, S. (2014). Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *The Journal of Neuroscience* 34, 817-822.
- Ito, M., and Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *The Journal of Neuroscience* 29, 9861-9874.
- Jedynak, J.P., Uslaner, J.M., Esteban, J.A., and Robinson, T.E. (2007). Methamphetamine-induced structural plasticity in the dorsal striatum. *European Journal of Neuroscience* 25, 847-853.
- Jentsch, J.D., Olausson, P., De La Garza, R., and Taylor, J.R. (2002). Impairments of reversal learning and response perseveration after repeated, intermittent cocaine administrations to monkeys. *Neuropsychopharmacology* 26, 183-190.
- Joel, D., and Weiner, I. (2000). The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. *Neuroscience* 96, 451-474.
- Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V., and Graybiel, A.M. (1999). Building neural representations of habits. *Science* 286, 1745-1749.
- Johnson, A., van der Meer, M.A.A., and Redish, A.D. (2007). Integrating hippocampus and striatum in decision-making. *Current Opinion in Neurobiology* 17, 692-697.

- Julian, M.D., Martin, A.B., Cuellar, B., Rodriguez De Fonseca, F., Navarro, M., Moratalla, R., and Garcia-Segura, L.M. (2003). Neuroanatomical relationship between type 1 cannabinoid receptors and dopaminergic systems in the rat basal ganglia. *Neuroscience* 119, 309-318.
- Jung, Y., Mulholland, P.J., Wiseman, S.L., Chandler, L.J., and Picciotto, M.R. (2013). Constitutive knockout of the membrane cytoskeleton protein beta adducin decreases mushroom spine density in the nucleus accumbens but does not prevent spine remodeling in response to cocaine. *The European Journal of Neuroscience* 37, 1-9.
- Kimchi, E.Y., and Laubach, M. (2009). Dynamic encoding of action selection by the medial striatum. *The Journal of Neuroscience* 29, 3148-3159.
- Komischke, B., Giurfa, M., Lachnit, H., and Malun, D. (2002). Successive olfactory reversal learning in honeybees. *Learning & Memory* 9, 122-129.
- Lee, D., Conroy, M.L., McGreevy, B.P., and Barraclough, D.J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cognitive Brain Research* 22, 45-58.
- Lichtman, A.H., and Martin, B.R. (1996). Δ^9 -tetrahydrocannabinol impairs spatial memory through a cannabinoid receptor mechanism. *Psychopharmacology* 126, 125-131.
- Lindgren, H.S., Wickens, R., Tait, D.S., Brown, V.J., and Dunnett, S.B. (2013). Lesions of the dorsomedial striatum impair formation of attentional set in rats. *Neuropharmacology* 71, 148-153.
- Lucantonio, F., Caprioli, D., and Schoenbaum, G. (2014). Transition from 'model-based' to 'model-free' behavioral control in addiction: Involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology* 76, Part B, 407-415.
- Martín, A.B., Fernandez-Espejo, E., Ferrer, B., Gorriti, M.A., Bilbao, A., Navarro, M., de Fonseca, F.R., and Moratalla, R. (2008). Expression and function of CB1 receptor in the rat striatum: localization and effects on D1 and D2 dopamine receptor-mediated motor behaviors. *Neuropsychopharmacology* 33, 1667-1679.

- Means, L.W., and Fernandez, T.J. (1992). Daily glucose injections facilitate performance of a win-stay water-escape working memory task in mice. *Behavioral Neuroscience* 106, 345.
- Miller, G.A. (1955). Note on the bias of information estimates. In: Quastler H, editor. *Information Theory in Psychology: Problems and Methods II–B*. Glencoe (Illinois): Free Press. pp. 95–100.
- Molina, L.A., Skelin, I., and Gruber, A.J. (2014). Acute NMDA receptor antagonism disrupts synchronization of action potential firing in rat prefrontal cortex. *PLoS ONE* 9, e85842.
- Montague, P.R., Dayan, P., and Sejnowski, T.J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience* 16, 1936-1947.
- Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., and Hikosaka, O. (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269-280.
- Nicola, S.M. (2010). The flexible approach hypothesis: Unification of effort and cue responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. *The Journal of Neuroscience* 30, 16585-16600.
- Niwa, M., Kamiya, A., Murai, R., Kubo, K.-I., Gruber, A.J., Tomita, K., Lu, L., Tomisato, S., Jaaro-Peled, H., Seshadri, S., Hiyama, H., Huang, B., Kohda, K., Noda, Y., O'Donnell, P., Nakajima, K., Sawa, A., and Nabeshima, T. (2010). Knockdown of DISC1 by in utero gene transfer disturbs postnatal dopaminergic maturation in the frontal cortex and leads to adult behavioral deficits. *Neuron* 65, 480-489.
- O'Doherty, J.P., Buchanan, T.W., Seymour, B., and Dolan, R.J. (2006). Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49, 157-166.
- Packard, M.G., and McGaugh, J.L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory* 65, 65-72.

- Pan, W.-X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *The Journal of Neuroscience* 25, 6235-6242.
- Parkinson, J., Dalley, J., Cardinal, R., Bamford, A., Fehner, B., Lachenal, G., Rudarakanchana, N., Halkerston, K., Robbins, T., and Everitt, B. (2002). Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behavioural Brain Research* 137, 149-163.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., and Frith, C.D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042-1045.
- Pontieri, F.E., Tanda, G., and Di Chiara, G. (1995). Intravenous cocaine, morphine, and amphetamine preferentially increase extracellular dopamine in the "shell" as compared with the "core" of the rat nucleus accumbens. *Proceedings of the National Academy of Sciences* 92, 12304-12308.
- Prendergast, B.J., Onishi, K.G., and Zucker, I. (2014). Female mice liberated for inclusion in neuroscience and biomedical research. *Neuroscience & Biobehavioral Reviews* 40, 1-5.
- Rayburn-Reeves, R.M., Laude, J.R., and Zentall, T.R. (2013). Pigeons show near-optimal win-stay/lose-shift performance on a simultaneous-discrimination, midsession reversal task with short intertrial intervals. *Behavioural Processes* 92, 65-70.
- Ritz, M.C., Lamb, R., and Kuhar, M. (1987). Cocaine receptors on dopamine transporters are related to self-administration of cocaine. *Science* 237, 1219-1223.
- Robinson, T.E., and Kolb, B. (1997). Persistent structural modifications in nucleus accumbens and prefrontal cortex neurons produced by previous experience with amphetamine. *The Journal of Neuroscience* 17, 8491-8497.
- Robinson, T.E., and Kolb, B. (1999). Alterations in the morphology of dendrites and dendritic spines in the nucleus accumbens and prefrontal cortex following repeated treatment with amphetamine or cocaine. *European Journal of Neuroscience* 11, 1598-1604.

- Robinson, T.E., and Kolb, B. (2004). Structural plasticity associated with exposure to drugs of abuse. *Neuropharmacology* 47, Supplement 1, 33-46.
- Roesch, M.R., Calu, D.J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience* 10, 1615-1624.
- Rolls, E.T., Loh, M., Deco, G., and Winterer, G. (2008). Computational models of schizophrenia and dopamine modulation in the prefrontal cortex. *Nature Reviews Neuroscience* 9, 696-709.
- Sañudo-Peña, M.C., Romero, J., Seale, G.E., Fernandez-Ruiz, J.J., and Walker, J.M. (2000). Activational role of cannabinoids on movement. *European Journal of Pharmacology* 391, 269-274.
- Sañudo-Peña, M.C., Tsou, K., and Walker, J.M. (1999). Motor actions of cannabinoids in the basal ganglia output nuclei. *Life Sciences* 65, 703-713.
- Schoenbaum, G., Nugent, S.L., Saddoris, M.P., and Setlow, B. (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* 13, 885-890.
- Schoenbaum, G., Saddoris, M.P., Ramus, S.J., Shaham, Y., and Setlow, B. (2004). Cocaine-experienced rats exhibit learning deficits in a task sensitive to orbitofrontal cortex lesions. *European Journal of Neuroscience* 19, 1997-2002.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology* 80, 1-27.
- Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593-1599.
- Seabold, G.K., Daunais, J.B., Rau, A., Grant, K.A., and Alvarez, V.A. (2010). DiOLISTIC Labeling of Neurons from Rodent and Non-human Primate Brain Slices. *Journal of Visualized Experiments : JoVE*, 2081.

- Seu, E., Lang, A., Rivera, R.J., and Jentsch, J.D. (2009). Inhibition of the norepinephrine transporter improves behavioral flexibility in rats and monkeys. *Psychopharmacology* 202, 505-519.
- Shen, H.-W., Toda, S., Moussawi, K., Bouknight, A., Zahm, D.S., and Kalivas, P.W. (2009). Altered dendritic spine plasticity in cocaine withdrawn rats. *The Journal of Neuroscience* 29, 2876-2884.
- Shen, H., Sesack, S., Toda, S., and Kalivas, P. (2008). Automated quantification of dendritic spine density and spine head diameter in medium spiny neurons of the nucleus accumbens. *Brain Structure and Function* 213, 149-157.
- Skelin, I., Hakstol, R., VanOyen, J., Mudiayi, D., Molina, L.A., Holec, V., Hong, N.S., Euston, D.R., McDonald, R.J., and Gruber, A.J. (2014). Lesions of dorsal striatum eliminate lose-switch responding but not mixed-response strategies in rats. *European Journal of Neuroscience* 39, 1655-1663.
- Sokolic, L., Long, L.E., Hunt, G.E., Arnold, J.C., and McGregor, I.S. (2011). Disruptive effects of the prototypical cannabinoid Δ^9 -tetrahydrocannabinol and the fatty acid amide inhibitor URB-597 on go/no-go auditory discrimination performance and olfactory reversal learning in rats. *Behavioural Pharmacology* 22, 191-202.
- Sorg, C., Manoliu, A., Neufang, S., Myers, N., Peters, H., Schwerthöffer, D., Scherr, M., Mühlau, M., Zimmer, C., Drzezga, A., Förstl, H., Bäuml, J., Eichele, T., Wohlschläger, A.M., and Riedl, V. (2013). Increased intrinsic brain activity in the striatum reflects symptom dimensions in schizophrenia. *Schizophrenia Bulletin* 39, 387-395.
- Staffend, N.A., and Meisel, R.L. (2011). DiOlistic labeling of neurons in tissue slices: a qualitative and quantitative analysis of methodological variations. *Frontiers in Neuroanatomy* 5, 14 eCollection 2011.
- Sulzer, D., Sonders, M.S., Poulsen, N.W., and Galli, A. (2005). Mechanisms of neurotransmitter release by amphetamines: a review. *Progress in Neurobiology* 75, 406-433.
- Svíženská, I., Dubový, P., and Šulcová, A. (2008). Cannabinoid receptors 1 and 2 (CB1 and CB2), their distribution, ligands and functional involvement in nervous system structures - A short review. *Pharmacology Biochemistry and Behavior* 90, 501-511.

- Tanda, G., Pontieri, F.E., and Di Chiara, G. (1997). Cannabinoid and heroin activation of mesolimbic dopamine transmission by a common μ 1 opioid receptor mechanism. *Science* 276, 2048-2050.
- Tepper, J.M., and Bolam, J.P. (2004). Functional diversity and specificity of neostriatal interneurons. *Current opinion in Neurobiology* 14, 685-692.
- Thapa, R., Wong, S., Sutherland, R.J., and Gruber, A.J. (Year). "Infusion of D2 receptor agonist in the dorsolateral striatum eliminate lose switch responding in rats", in: *Society for Neuroscience Annual Meeting*.
- Thorn, C.A., Atallah, H., Howe, M., and Graybiel, A.M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* 66, 781-795.
- Van Waes, V., Beverley, J., Siman, H., Tseng, K.Y., and Steiner, H. (2012). CB1 cannabinoid receptor expression in the striatum: Association with corticostriatal circuits and developmental regulation. *Frontiers in Pharmacology* 3.
- Varvel, S., Hamm, R., Martin, B., and Lichtman, A. (2001). Differential effects of Δ 9-THC on spatial reference and working memory in mice. *Psychopharmacology* 157, 142-150.
- Voorn, P., Vanderschuren, L.J., Groenewegen, H.J., Robbins, T.W., and Pennartz, C.M. (2004). Putting a spin on the dorsal-ventral divide of the striatum. *Trends in Neurosciences* 27, 468-474.
- Wang, Z., Xu, B., and Zhou, H.-J. (2014). Social cycling and conditional responses in the Rock-Paper-Scissors game. *Scientific Reports*. 4.
- Weinstein, A., Brickner, O., Lerman, H., Greemland, M., Bloch, M., Lester, H., Chisin, R., Sarne, Y., Mechoulam, R., and Bar-Hamburger, R. (2008). A study investigating the acute dose—response effects of 13 mg and 17 mg Δ 9-tetrahydrocannabinol on cognitive—motor skills, subjective and autonomic measures in regular users of marijuana. *Journal of Psychopharmacology* 22, 441-451.
- Wilkinson, L.S., Mittleman, G., Torres, E., Humby, T., Hall, F.S., and Robbins, T.W. (1993). Enhancement of amphetamine-induced locomotor activity and dopamine

release in nucleus accumbens following excitotoxic lesions of the hippocampus. *Behavioural Brain Research* 55, 143-150.

Willuhn, I., Burgeno, L.M., Everitt, B.J., and Phillips, P.E.M. (2012). Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. *Proceedings of the National Academy of Sciences* 109, 20703-20708.

Wong, S.A., Thapa, R., Badenhorst, C.A., Briggs, A.R., Sawada, J.A., and Gruber, A.J. (2016). Opposing effects of acute and chronic d-amphetamine on decision-making in rats. *Neuroscience* S0306-4522(16)30086-0.

Wright, M., Vandewater, S.A., Parsons, L.H., and Taffe, M.A. (2013). $\Delta 9$ Tetrahydrocannabinol impairs reversal learning but not extra-dimensional shifts in rhesus macaques. *Neuroscience* 235, 51-58.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience* 19, 181-189.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2005a). Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *European Journal of Neuroscience* 22, 505-512.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005b). The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience* 22, 513-523.

Zeeb, F.D., Robbins, T.W., and Winstanley, C.A. (2009). Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. *Neuropsychopharmacology* 34, 2329-2343.