DISSECTING THE DETERMINANTS OF DECISIONS: SEX DIFFERENCES AND DISTRIBUTED REINFORCEMENT REVEALED IN A RAT CHOICE TASK

CLIFFORD DONOVAN
Bachelor of Science, University of Victoria, 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

DISSECTING THE DETERMINANTS OF DECISIONS: SEX DIFFERENCES AND DISTRIBUTED REINFORCEMENT REVEALED IN A RAT CHOICE TASK

CLIFFORD DONOVAN

Date of Defence: April 4th, 2019

Dr. A. Gruber Supervisor	Associate Professor	Ph.D.
Dr. R. Gibb Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. D. Euston Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. M. Tata Chair, Thesis Examination Committee	Associate Professor	Ph.D.

Dedication

This work is dedicated to my wife Miranda, who helped to maintain my mental health throughout the preparation of this thesis. Without her support, this work would not have been completed.

Abstract

Decision-making in the mammalian brain involves structures within the midbrain, striatum, limbic system, and cortex. My colleagues and I studied the roles of several of these structures in reward-processing and decision-making related phenomena. First, we found female rats were more likely to approach feeders outside the task context in an operant chamber with two feeders. I speculate this sexual dimorphism relates to the disproportionate effect of fasting on reproduction in female rats; this provides an alternative to the widespread interpretation that male rats are more risk-seeking than females. Next, we recorded field potentials (FPs) from seven brain regions as rats completed a binary choice task to determine if reward information originated in a particular area, and if the fidelity of information varied among them. Using a machine learning classifier, we found reinforcement information was distributed across the network and there was no canonical flow of information in the recorded structures.

Acknowledgements

I sincerely thank the members of Dr. Gruber and Dr. Gibb's labs that completed all of the experimental work for these projects; without them there would have been no data for my analysis. I especially thank Cecilia Badenhorst for her electrophysiology work (chapter 4). I also thank Surjeet Singh, Ali Mashhoori, Saeedeh Hashemnia, Sienna Randolph, Scott Wong, Vicky Ivan, Chelsea Matisz, Dr. David Euston, and Dr. Masami Tatsuno, for their aid in overcoming obstacles in analysis and editing of the text.

I also thank my supervisor and mentor, Dr. Aaron Gruber. He listened to my rants, both political and otherwise; he listened to my ridiculous theories of brain function; and he listened to me lose faith in my project many times. Most importantly, Dr. Gruber gave me this tremendous opportunity to come from a background of no programming expertise and learn everything necessary to do some amazing, state of the art analysis using deep learning. The skills he helped me acquire reach far beyond academia, and I am forever grateful that he believed in my curiosity.

Table of Contents

List of Figures

List of Abbreviations

ACC	anterior cingulate cortex
ANOVA	analysis of variance
AP	anterior - posterior
BLA	basolateral amygdala
DA	dopamine
DLS	dorsolateral striatum
DMS	dorsomedial striatum
DV	dorsal – ventral
EEG	electroencephalogram
EFS	extraneous feeder sampling
FIR	finite impulse response
fMRI	functional magnetic resonance imaging
FP	field potential
HPC	hippocampus
IGT	Iowa gambling task
ITI	inter-trial interval
LE	Long Evans
mPFC	medial prefrontal cortex
ML	medial - lateral
NACc	nucleus accumbens core
PFC	prefrontal cortex
RDT	risky decision making task
ReLU	rectified linear unit
SEM	standard error of the mean
OFC	orbitofrontal cortex
VST	ventral striatum

1. General Introduction

Our brains make an astounding amount of decisions throughout each day, most without conscious thought. For example, in the morning commute to work, a decision is made of which route to take, how quickly to drive, and to follow the traffic laws. Underlying elements that allow this commute also involve decisions that do not engage attention, such as which hand to use to open the car door, or to not roll the windows down because it is cold outside. The hierarchical complexity of these, and other decision-making behaviours, unsurprisingly involve a large proportion of the human brain (Vickery, Chun, & Lee, 2011). Consequently, a plethora of psychological disorders are associated with dysfunctional components of this system (Brunello, Masotto, Steardo, Markstein, & Racagni, 1995; Dauer & Przedborski, 2003; Everitt & Robbins, 2005; Solanto, 1998).

The decision-making system comprises an intricate neural circuit involving cortical, striatal, limbic, thalamic, and midbrain structures (Burton, Nakamura, & Roesch, 2015; Gruber & McDonald, 2012; Haber, 2003; Voorn, Vanderschuren, Groenewegen, Robbins, & Pennartz, 2004). The hippocampus (HPC) seems to process information related to the state of the animal and its current environment, as well as integrate information to adhere to a specific temporal sequence (Euston, Gruber, & McNaughton, 2012; Gruber & McDonald, 2012; Pennartz, Ito, Verschure, Battaglia, & Robbins, 2011). Lesion and inactivation studies suggest the basolateral nucleus of the amygdala (BLA) is involved in attentional control and stimulus-outcome learning (Roesch, Calu, Esber, & Schoenbaum, 2010). Specifically, the BLA seems to encode the salience of a stimuli and relay this information to the cortex and striatum (Ambroggi, Ishikawa, Fields, & Nicola,

2008; Schoenbaum, Setlow, Saddoris, & Gallagher, 2003). The medial prefrontal cortex (mPFC) appears to be involved in contextual associations (Barker, Bird, Alexander, & Warburton, 2007), the temporal representation of reward contingency (Coutureau, Esclassan, Di Scala, & Marchand, 2012), the comparison of cost and expected value of choices (Shenhav, Cohen, & Botvinick, 2016), and the assessment of counterfactual outcomes (Mashhoori, Hashemnia, McNaughton, Euston, & Gruber, 2018). These functions appear to integrate reinforcement information, whether real or expected, to form policies of responding, which inform the mPFC in its top-down control over motor actions (Euston et al., 2012). Finally, the orbitofrontal cortex (OFC) seems to be necessary for reversal learning, and encoding the expected value of outcomes (Murray, O'Doherty, & Schoenbaum, 2007; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2009).

The striatum, which, in addition to thalamic and midbrain input, receives projections from all of the aforementioned brain regions, appears to have distinct functional subdivisions, following a dorsolateral to ventromedial divide (Voorn et al., 2004). The ventral striatum (VST), comprising the nucleus accumbens core and shell, is commonly implicated in cue-associations and conditioning with reinforcement (Burton et al., 2015; Ito & Doya, 2009). This striatal subdivision seems to combine affective, spatial, and contextual information to develop an affective valence by which it influences behaviour through output to the hypothalamus and reciprocal projections to dopaminergic (DA) centres (Floresco, St. Onge, Ghods-Sharifi, & Winstanley, 2008; Gruber, Hussain, & O'Donnell, 2009; Voorn et al., 2004). The dorsomedial striatum (DMS) is implicated in learning response-outcome (instrumental) associations (Burton et al., 2015; Gruber & McDonald, 2012). Specifically, the DMS has been shown to encode the relative value

between possible actions (Wang, Miura, & Uchida, 2013), including an action-specific reward prediction error, which has not been observed in the ventral or dorsolateral subdivisions of the structure (Roesch, Singh, Brown, Mullins, & Schoenbaum, 2009; Stalnaker, Calhoon, Ogawa, Roesch, & Schoenbaum, 2012). Lastly, the dorsolateral striatum (DLS) receives somatotopically organized connections from sensorimotor areas of the cortex, and is thus more directly involved in motor action selection and stimulus-response associations (Voorn et al., 2004). As a stimulus is repeatedly reinforced, by dopaminergic modulation, a more habit-like sensorimotor association begins to form, in which the DLS is suggested to be more involved (Gruber & McDonald, 2012).

All of these aforementioned structures of the cortico-striatal-limbic circuit are influenced by dopamine (Björklund & Dunnett, 2007; Bromberg-Martin, Matsumoto, & Hikosaka, 2010; Hitchcott, Quinn, & Taylor, 2007; Packard & White, 1991; Phillips, Salussolia, & Hitchcott, 2010). These neurons primarily originate in the midbrain structures of the substantia nigra and ventral tegmental area (Björklund & Dunnett, 2007). A prevalent theory suggests two different types of dopamine neurons: some encoding a prediction error signal (Berridge & Robinson, 1998; Schultz, Dayan, & Montague, 1997), and others encoding a more general salience of stimuli regardless of valence (Bromberg-Martin et al., 2010; Horvitz, 2000). Although we do not investigate the role of dopamine directly, we suspect the former function—encoding affective valence and motivational value of reinforcement—is a prevalent driver of the results I detail in the following experiments.

This thesis explores decision-making with two separate experiments, both utilizing a binary choice task. First, we investigated possible sex differences in multiple

measures on this behavioural task. We found female rats were more likely to extraneously sample the unchosen feeder outside of the task context. This behaviour caused an errant conclusion of sex differences within the constraints of the task. Controlling for this extraneous behaviour eliminated this sex difference. This odd behaviour, and its effect on the measures in our task, elucidated our ignorance to more fundamental aspects of choice, such as immediate processing at the time of reinforcement. Thus, we recorded field potential activity from the seven previously mentioned cortico-striatal-limbic regions as rats completed the same choice task. We found outcome valence, immediately following the reinforcement event, was encoded in all of the recorded regions. However, no single brain structure contained all of the information available within the remainder of the network.

2. General Methods

2.1 Animal Housing

Rats were pair-housed in plastic cages with corncob bedding and a piece of PVC pipe for enrichment. On behavioural testing days, animals were restricted to one hour of water, but otherwise had *ad libitum* access to food and water. All testing and procedures were approved by the University of Lethbridge Animal Welfare Committee and comply with the Canadian Council on Animal Care.

2.2 Behavioural Task

Behavioural testing was performed in aluminum operant chambers (Fig. 1; Skelin et al., 2014). Rats were placed in the operant chamber for 45 minute sessions. Trials were self-paced, and initiated by the rat performing a nose-poke into the central port. Following nose-poke entry (>150ms duration), a tone (6 KHz) was presented to indicate that the animal could then locomote to one of the two adjacent sucrose delivery feeders. Entry to the feeder well was detected by infrared emitters and sensors. If the animal chose the correct well, reward (60 µL of 10% sucrose solution) was immediately delivered by activation of a solenoid valve. 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-light turned off; panel light turned on). This change in lighting served to indicate that reward was not forthcoming, and was of sufficiently short duration such that it terminated by the time the rats returned to the central poke port; there was therefore no 'time-out' associated with reward omission. Once a feeder was chosen, or if no feeder was chosen in the 15 s following a nose-poke, the trial ended and the rat had to return to the central port to initiate a new trial.

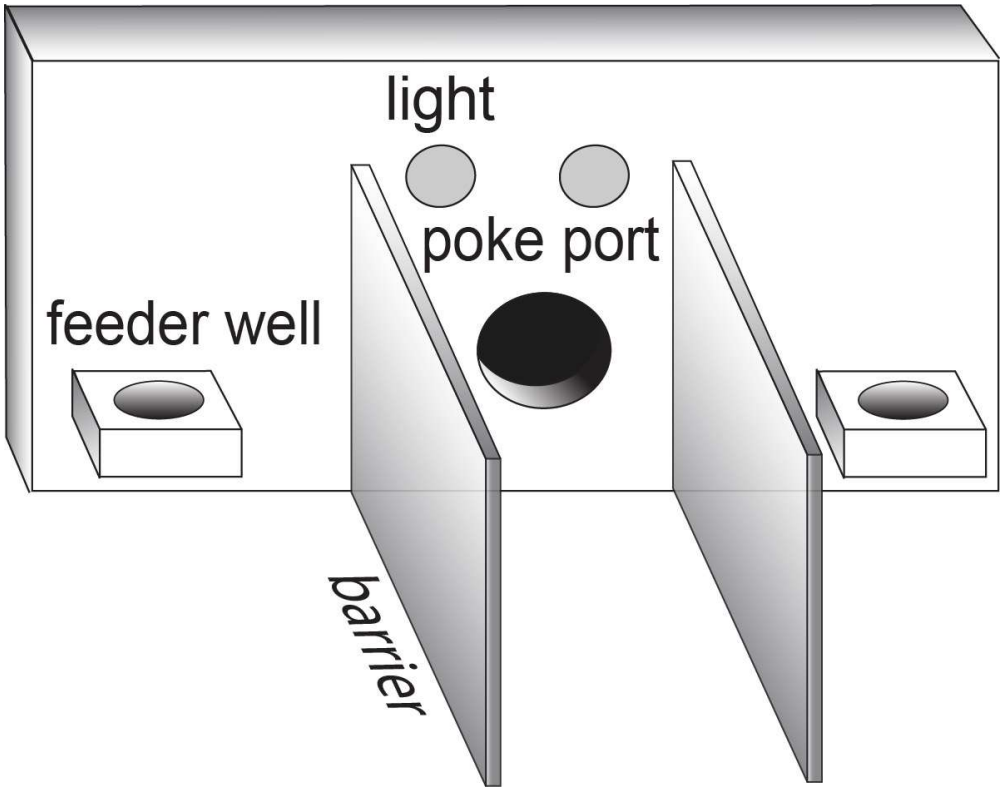

Figure 1. Illustration of the behavioural apparatus. The two panel lights illuminate, and the overhead house light extinguishes to indicate the rat is able to begin a trial. To initiate a trial, the rat pokes its snout into the center port. The rat then traverses around the barrier (13.5cm in length) to a feeder well.

The behaviour of animals was shaped during the first two training sessions. In the first session, all trials were rewarded to facilitate task acquisition. In the second training session, reward probability was reduced to 0.5. Following these sessions, reinforcement was controlled by an algorithm that attempted to minimize the number of rewards given to the rats by predicting which feeder it would select (Lee, Conroy, McGreevy, & Barraclough, 2004; Skelin et al., 2014). This was done by examining the choices and reinforcements from the previous four trials. If either feeder was selected at a greater than chance rate in the context of these past trials, it would be unrewarded for the upcoming trial. In doing so, the competitive mode implements the classic 'Matching Pennies' task. Optimal performance (random feeder well selection) would result in reward on 50% of

the trials. 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, 13.5cm) 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. Sex differences in rat decision-making: The confounding role of extraneous feeder sampling between trials

3.1 Introduction

Men and women sometimes differ in the way they use past rewards to guide future choices. It has been suggested that men are more likely to exhibit risk-taking behaviour than women (Becker, Perry, & Westenbroek, 2012; Byrnes, Miller, & Schafer, 1999; Cross, Copping, & Campbell, 2011), whereas women have been suggested to be more sensitive to loss than men (Cross et al., 2011; Van den Bos, Homberg, & de Visser, 2013). Much of the supporting evidence for these sex differences comes from tasks in which subjects choose among options with different expected values, the most prominent of which is the Iowa Gambling Task (IGT). There is strong evidence that men develop an undeviating preference for the optimal choice in fewer trials than do women (for review see: Van den Bos et al., 2013). This difference in strategy has been interpreted as women exhibiting heightened loss-sensitivity relative to men. This interpretation is supported by a recent meta-analysis of several other decision-making tasks (Cross et al., 2011).

Rodent studies of decision-making have revealed similar disparities due to sex in some situations (Jentsch & Taylor, 2003; Orsini, Willis, Gilbert, Bizon, & Setlow, 2016), but the evidence is far less conclusive (for review see: Orsini & Setlow, 2017). In a rodent analogue of the IGT, male Wistar rats collected more reward than females (Van den Bos, Jolles, Van der Knaap, Baars, & De Visser, 2012). However, the same investigators found no sex differences when testing Long Evans rats on the same task (van Hasselt et al., 2012). Using a different adaptation of the IGT for rodents (Zeeb, Robbins, & Winstanley, 2009), another research group found no sex-based differences in Sprague-Dawley rats

(Peak, Turner, & Burne, 2015). Other tasks have been utilized to investigate additional facets of rat decision-making, such as the risky decision-making task (RDT). In the RDT, rats choose between a safe lever, in which they consistently receive a small food reward, and a risky lever, in which they receive a larger food reward accompanied by an increasingly higher chance of receiving a foot shock. Male Long Evans rats chose the risky lever significantly more than the females (Orsini et al., 2016). Similar to results from human subjects, this effect may be interpreted as a measure of heightened loss-sensitivity in females or heightened risk-taking behaviour in males. Male Sprague-Dawley rats also displayed more impulsive responding than their female counterparts on a signal discrimination task (Jentsch & Taylor, 2003). However, contrary results have been found using delayed discounting tasks, which, are a direct measure of impulsive choice. In this paradigm, animals choose between a small, immediate reward and a larger, delayed reward. There has been no sex differences suggested from studies utilizing delayed discounting tasks in several strains of adult, drug naïve rats, including Long Evans rats (Eubig, Noe, Floresco, Sable, & Schantz, 2014), Sprague Dawley rats (Lukkes, Thompson, Freund, & Andersen, 2016), or Wistar rats (Smethells, Swalve, Eberly, & Carroll, 2016).

The inconsistency in the rat literature raises questions about the generalization of sex discrepancies in the choice domain across mammalian brains. It is possible that this inconsistency is the product of some unexplained factor that is confounding the results. The control of motivated behaviour is the product of interactions among several brain networks that process information in unique ways (Balleine & O'Doherty, 2010; Gruber & McDonald, 2012). Choice behaviour by rats and humans is often better explained by

taking into account such interactions (Devan, Hong, & McDonald, 2011; Ito & Doya, 2009; Skelin et al., 2014). Examining the effect of biological sex on specific behaviours mediated by these distinct brain systems may help explain the apparent inconsistency of past reports. One specific behaviour is Pavlovian approach, which fosters orientation and approach toward rewarding stimuli, such as feeders in an experimental chamber. This is an intrinsic behaviour that can affect choice. For instance, rats will approach nearby feeders more often than distant ones, even if the nearby feeder delivers suboptimal reward (Morrison & Nicola, 2014). Moreover, these approaches affect subsequent choices (Gruber, Thapa, & Randolph, 2017; Wong, Thapa, et al., 2017). Pavlovian approach is ubiquitous across tasks, and is subject to significant variation among individuals (Flagel, Watson, Robinson, & Akil, 2007; Kearns, Gomez-Serrano, Weiss, & Riley, 2006; Pitchers et al., 2015). Further, sex differences have also been observed in Pavlovian approach (Hammerslag & Gulley, 2014; Pitchers et al., 2015). Thus, it is possible that sex differences in performance on decision-making tasks may be confounded with Pavlovian approach. Moreover, factors such as apparatus design or strain may influence such approach (Meyer et al., 2012; Robinson, Yager, Cogan, & Saunders, 2014), and thereby indirectly affect choice to a larger degree in one sex.

Here we used a well-validated task with unpredictable rewards in order to decompose reinforcement-driven shifts in decisions into several components (Gruber & Thapa, 2016; Skelin et al., 2014; Wong, Thapa, et al., 2017). In our task, we are able to assess sensitivity to wins or losses, motivation, and feeder approach behaviour. Specifically, we examine the relationship between motoric measures and choice strategies. These strategies include: ‘lose-shift’ responding (the animal’s propensity to

alter their responding after reward absence/punishment); 'win-stay' responding (the animal's likelihood to repeat an action upon receipt of reward); and a newly reported approach behaviour we call extraneous feeder sampling (EFS), wherein rats sample the alternate feeder prior to initiating the subsequent trial (Gruber et al., 2017; Wong, Randolph, Ivan, & Gruber, 2017; Wong, Thapa, et al., 2017). In light of past research from other labs summarized above (Jentsch & Taylor, 2003; Orsini & Setlow, 2017; Orsini et al., 2016; Van den Bos et al., 2013, 2012), we expected females to exhibit increased loss-sensitivity as compared to males. However, loss-sensitivity provides an imprecise connotation. Sensitivity to loss may refer to an emotional frustration, a devaluation of reward in a reinforcement learning context, immediate motor behaviour following reward omission, or other responses. In our task, we are specifically referring to the lose-shift response: the immediate decision of the animal to shift feeder choice following reward omission. This appears to be distinct from forms of reinforcement learning that integrate information over several trials (Gruber & Thapa, 2016). The data presented here suggest that sex-based differences in lose-shift reinforcement are weak or nonexistent, but that there is a difference in feeder approach between trials that can induce an apparent effect of loss sensitivity if not properly controlled. Between-trial behaviour should thus be taken into account so as to avoid misattributing differences in feeder approach to differences in risk, loss-sensitivity, or other factors influencing choice.

3.2. Methods

3.2.1. Animals and Experimental Design

We collected behavioural performance data from 106 rats in three separate cohorts. Each cohort contained both male and female animals. All animals were born in

our facility, were housed under the same conditions, and were trained using the same protocol. Animals that did not complete at least 150 trials in the testing session were removed from analysis. This exclusion criteria left us with data from three cohorts consisting of: Cohort 1: 28 Long Evans (15 male, 13 female, 71-117 days old); Cohort 2: 23 Long Evans rats (17 male, 6 female, 80-103 days old); and Cohort 3: 28 Long Evans rats expressing a transgene in some cells (Cre+; 13 male, 15 female; 71-112 days old). The animals from Cohort 3 expressed a transgene (cre-recombinase) under the control of the Tyrosine Hydroxylase promoter (see Witten et al., 2011 for more details), but had no other manipulations. These animals were included to ascertain whether these germline genetic manipulations to dopamine neurons had a baseline effect on decision-making. While the Cre-lox system is widely used in controlling transcription and translation of specific cell populations, recent studies have called into question the potential for cre toxicity (Buerger et al., 2006; Forni et al., 2006), DNA damage (Loonstra et al., 2001), and illegitimate chromosome rearrangement (Schmidt, Taylor, Prigge, Barnett, & Capecchi, 2000) with the use of these genetic tools. This transgenic cohort was not statistically different from the others, so their data were pooled with the other cohorts, giving a total of 79 animals (45 male, 34 female) in the study.

In order to account for possible differences in learning over days, we have only taken data from the fourth session in which the long barriers (13.5cm) were used. These data comprised a total of 19,073 trials.

3.2.2. Statistical analysis

Eight variables were measured on the task. These measures can be separated into motor components and decision-making components. Motor components consisted of: the

number of trials completed within the session; inter-trial interval (ITI, time between leaving a feeder and performing a nosepoke to initiate a new trial); and response time (time to locomote from central nosepoke hole to either adjacent feeder). The decision-making components were calculated over the entire session length and consisted of: rewarded trials (the percentage of correct feeder choices the animal made); win-stay (the probability of returning to the same feeder that provided a reward on the previous trial); lose-shift (the probability of choosing the opposite feeder after not being rewarded on the previous trial); and extraneous feeder sampling (the probability of the animal to sample both wells within one trial; EFS; Wong et al., 2017). If a rat sampled both feeders (EFS), the 'shift' or 'stay' is computed with respect to the first feeder sampled following the nosepoke (i.e. the choice). EFS is calculated as a probability to exhibit this behaviour, taken from the total number of trials throughout the session; thus, it is a relative measure of this behaviour and is invariant to differences due to locomotor activity. This multi-feeder sampling is never rewarded, yet persists over thousands of trials and does not extinguish (Gruber & Thapa, 2016). This behaviour has led us to speculate that EFS is not a lack of task comprehension. All statistical analyses were performed using SPSS version 22 (IBM Corp., New York) and MATLAB version R2015a (The Mathworks Inc., Massachusetts).

We first analyzed the potential effects of the transgene on measures in our task using a one-way Analysis of Variance (ANOVA). Finding no significant differences, the data was then pooled across all the animals and two-tailed t-tests ($\alpha = 0.05$) were performed to compare sex differences in session-averaged behavioural measures. When measures exhibited unequal variance according to Levene's test, ($\alpha = 0.05$), Welch's t-test

was used with the Welch-Satterthwaite equation to approximate the degrees of freedom (Hall & Willink, 2001).

To compare the change in behaviour within a session, data was binned into eight bins of five minutes each to account for 40 minutes of the 45 minute session. The final five minutes of each session were not included to avoid the potential confound of animals being distracted by the experimenter returning to the room. Of the 79 animals included in the study, two did not complete at least a 40 minute session and were excluded from this within-session analysis. The bins were then plotted against each of the task variables using MATLAB. A mixed model ANOVA was performed to compare within group (time bin) and between group (sex) variables. The Greenhouse-Geisser correction was applied to all the measures on the mixed model ANOVA, as they violated the assumption of sphericity, as tested by Mauchly's test of sphericity ($\alpha = 0.05$).

3.3. Results

We found that female rats engaged in extraneous feeder sampling (EFS) more often than males ($t = 2.60$, Welch-Satterthwaite $df = 52.6$, $p = 0.012$, $d = 0.612$; Fig 2A). Moreover, EFS was higher in females regardless if rats received a reward in the trial ($t = 2.55$, Welch-Satterthwaite $df = 53.1$, $p = 0.014$, $d = 0.601$; Fig. 2B) or not ($t = 2.37$, $df = 77$, $p = 0.021$, $d = 0.530$; Fig. 2C). Therefore, female rats approach the feeder outside of the task sequence more than males, regardless of reward outcome.

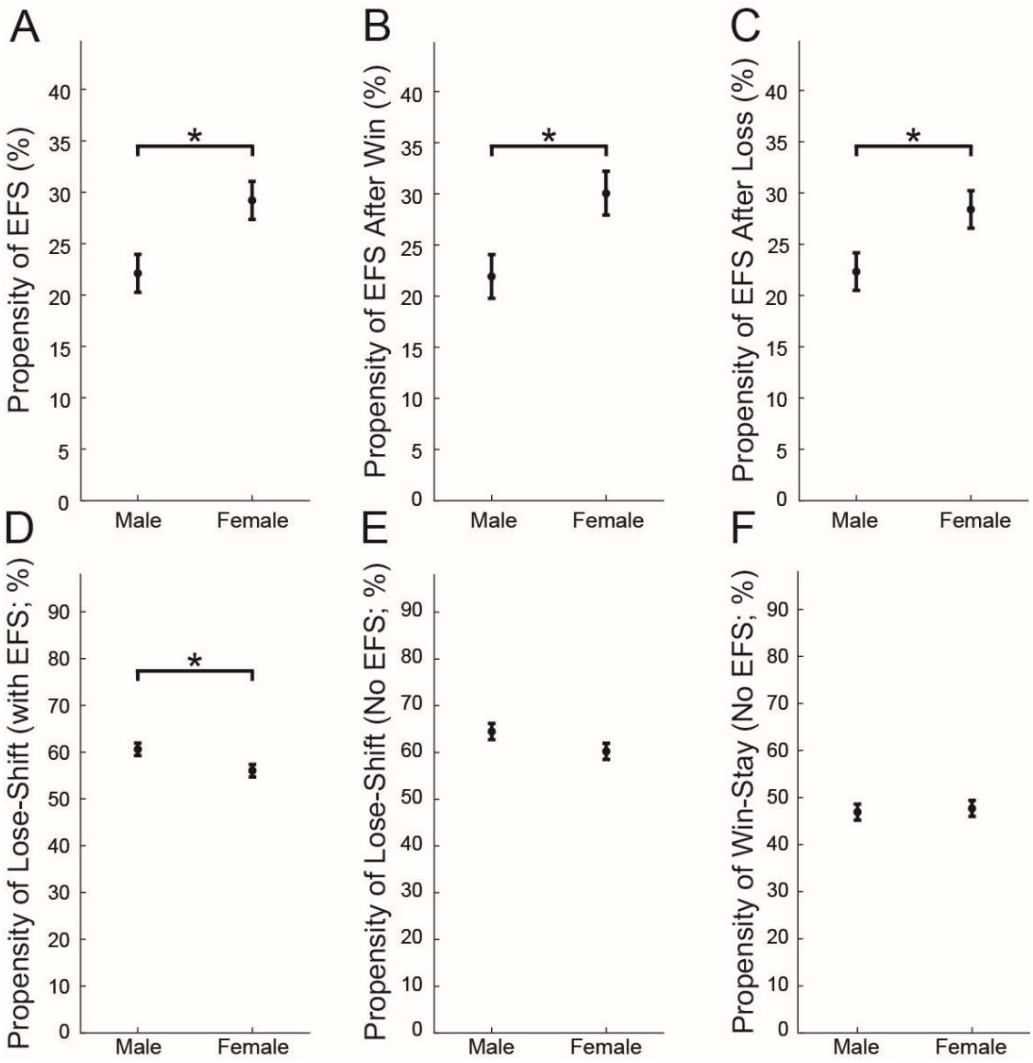

Figure 2. Sex differences in behaviour. (A) The propensity of rats to sample both wells within one trial (EFS) was significantly greater in females. (B) The propensity of rats to sample both wells within one trial following reward was significantly greater in females. (C) The propensity of rats to sample both wells within one trial following reward omission was also significantly greater in females. (D) When trials following EFS were included, males were significantly more likely to switch feeder choice following reward omission. (E) When trials following EFS were excluded, the percentage of trials in which the rats switched feeder choice after a loss did not differ significantly between sexes, but a trend for males to lose-shift more remained. (F) The percentage of trials, excluding those following EFS, in which the rats maintained feeder choice after a win did not vary between sexes. Error bars indicate standard error of the mean and asterisks '*' indicate statistically different means as determined by the two-tailed t-test ($p < 0.05$).

Because female rats generate more EFS, and the EFS could affect subsequent choice (Gruber et al., 2017), we expected some sex-based differences in choice to arise as

a consequence of different EFS rates. We therefore quantified the propensity to shift responses following either reinforcement delivery (win) or reinforcement omission (loss). Indeed, analysis of lose-shift across all trials (including those after EFS) showed that males were more likely to lose-shift than females (lose-shift: $t = 2.16$, $df = 77$, $p = 0.034$, $d = 0.548$; Fig. 2D). There was no significant difference between the sexes in their probability to win-stay ($t = 0.103$, $df = 77$, $p = 0.918$, $d = 0.035$). However, to further test for a confounding role of EFS, we tested for sex-based differences exclusive of EFS by analyzing only those trials that did not follow EFS. This exclusion resulted in rejection of 4,669 out of 19,073 trials. By eliminating these EFS-preceded trials, we found no significant sex difference in lose-shift ($t = 1.77$, $df = 77$, $p = 0.081$, $d = 0.397$; Fig. 2E) or win-stay ($t = 0.302$, $df = 77$, $p = 0.764$, $d = 0.070$; Fig. 2F). To ensure this change in statistical significance in lose-shift was not due to the reduction in sample size, we randomly removed 4669 and repeated the lose shift comparison. This resulted in a similar statistical significance to the original result when EFS trials were included ($t = 2.13$, $df = 77$, $p = 0.037$). Thus, the observed increased EFS in females was exerting a confounding effect and causing an apparent decrease in lose-shift responding. This is likely caused by the calculation of lose-shift from the original well sampled. Thus, the animals were completing a lose-shift from the second sampled well, which would be recorded as a lose-stay response. Decreased lose-shift responding should improve performance on the present task because it is a less predictable strategy. However, we saw no significant sex difference in the number of rewarded trials when either including ($t = 0.931$, $df = 77$, $p = 0.355$, $d = 0.214$) or excluding ($t = 1.78$, $df = 77$, $p = 0.080$, $d = 0.402$) trials following

EFS. This is likely because the competitive algorithm is not strong enough to fully penalize moderate levels of lose-shift responding.

The probability of lose-shift and win-stay responding on this task depend on the inter-trial interval (ITI, the duration between reinforcement and the subsequent trial). Specifically, we have previously shown probability to lose-shift follows a log-linear negative relationship with increasing ITI, reaching chance levels beyond 7 seconds. Whereas win-stay follows a log-parabolic relationship with ITI, with the highest probability to win-stay at approximately an 8 second ITI (Gruber & Thapa, 2016). This indicates that the speed of the animal to complete trials should have an effect on the likelihood of shifting choice after a loss or a win. Excluding trials following EFS, the mean ITI following wins ($t = 1.62$, $df = 77$, $p = 0.101$, $d = 0.355$; Fig. 3A) or losses ($t = 0.849$, $df = 77$, $p = 0.398$, $d = 0.190$; Fig. 3B) was not different between sexes. Interestingly, there was a non-significant trend for the male rats to be slower in going from the nose-poke to the reward feeders ($t = 1.83$, $df = 77$, $p = 0.070$, $d = 0.414$; Fig. 3C). These data indicate that there may be differences in movement speed on the task, which could affect choice. We suspect this difference is likely due to differences in body size and weight, rather than a difference in motivational drive.

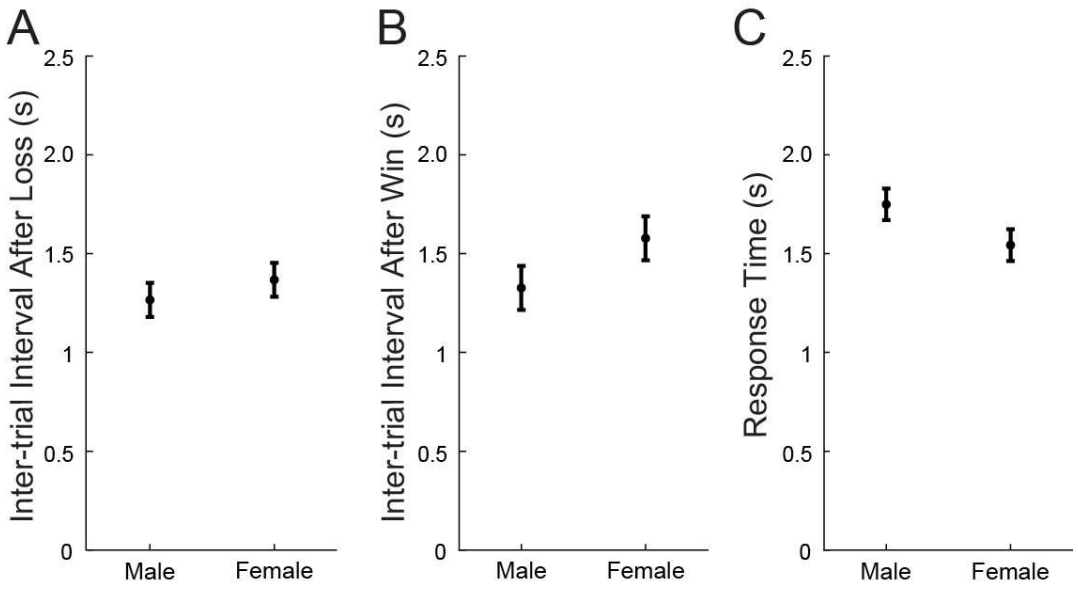

Figure 3. Sex differences in motivation and motoric speed on the choice task. (A) The time interval from loss reinforcement to the next nose-poke did not differ between sexes. (B) The time to start a new trial following a win did not vary between sexes. (C) There was a non-significant trend for female rats to be faster than males in their locomotion to the feeder well following trial initiation. None of the tested means were significantly different, as determined by the two-tailed t-test ($p < 0.05$). Error bars indicate standard error of the mean.

These sex-based disparities appear to be primary differences in decision-making and not artifacts of performance or motivation. However, motivation does change within the session as animals become sated. Because males and females differ in weight and calorie consumption (Wade, 1972), it could be that their motivation changes differently during the session. For instance, females could become sated more quickly and therefore become less sensitive to reward omission as the session progresses. We investigated this by quantifying the dependent response variables in bins of time during the session. The session was broken into eight time bins of five minutes each and a mixed model (repeated measures) ANOVA was used to test for statistical significance of the means. We found that females performed significantly more EFS throughout the session (Main effect: $F(1, 75) = 7.83$, $p = 0.007$, $\eta^2 = 0.095$; Fig. 4A). In order to eliminate the possible confounding

role of EFS on other response variables, we eliminated the trials following EFS for subsequent analysis. There was no significant main effect due to sex on the number of trials completed ($F(1, 75) = 1.342$, $p = 0.250$, $\eta^2 = 0.018$; Fig. 4B), or the number of rewarded trials over the time bins ($F(1, 75) = 3.01$, $p = 0.087$, $\eta^2 = 0.039$; Fig. 4C). These data suggest that motivation is not different between sexes within a session. There also was no main effect of sex on the probability of lose-shift ($F(1, 76) = 2.04$, $p = 0.157$, $\eta^2 = 0.026$; Fig. 4D), or the probability of win-stay ($F(1, 76) = 0.040$, $p = 0.843$, $\eta^2 = 0.001$) over all time bins.

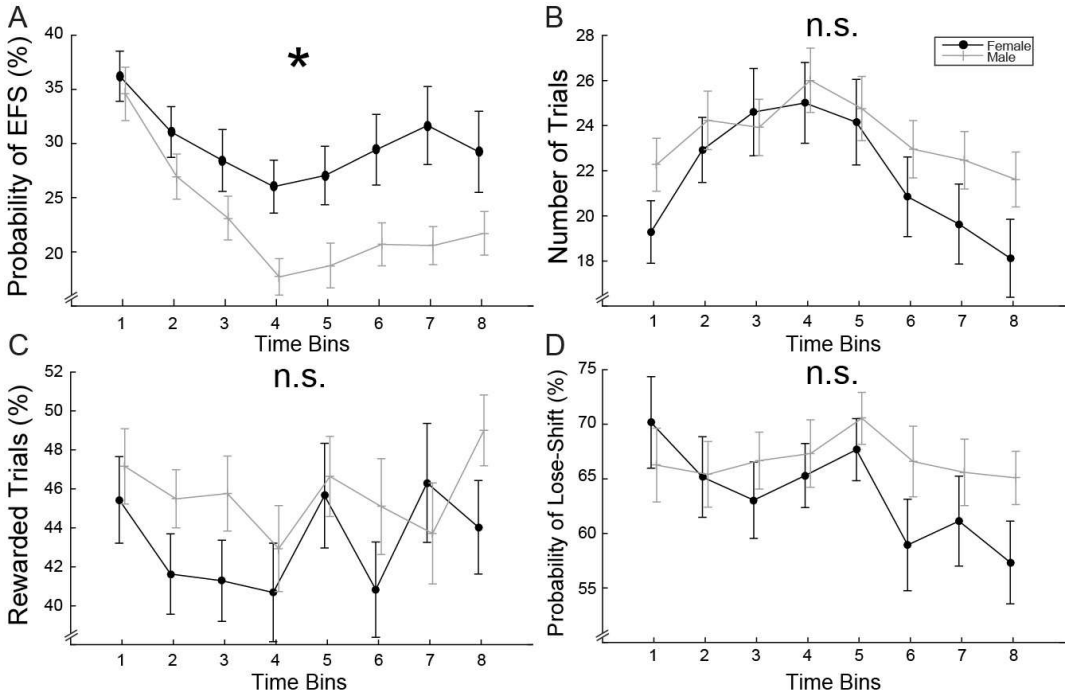

Figure 4. Sex differences in within-session task performance. (A) The probability to sample both feeders between trials decreased as sessions progressed, but females exhibited a higher rate of this behaviour throughout the session. (B) There was no significant difference in the number of trials completed over the time bins between the sexes. (C) There was no significant difference between the sexes on the rats' percentage of rewarded trials during the session. (D) There was no significant difference between the sexes on the rats' probability to shift feeder choice after a loss. Trials following the rats sampling both wells were excluded in panels B-D. '*' indicates a significant main effect of sex by mixed model ANOVA ($p < 0.05$); 'n.s.' indicates no significance. Error bars indicate standard error of the mean.

There was no significant interaction effect between sex and time bins on any of the dependent response variables. The Greenhouse-Geisser correction was used for calculation of these statistics as they violated the assumption of sphericity by Mauchly's test. These variables include the number of trials completed ($F(4.86, 365) = 0.970$, $p = 0.435$, $\eta^2 = 0.013$), the number of rewarded trials ($F(5.79, 434) = 0.562$, $p = 0.755$, $\eta^2 = 0.007$), probability of EFS ($F(4.50, 338) = 1.34$, $p = 0.250$, $\eta^2 = 0.018$), probability of lose-shift ($F(5.93, 445) = 0.556$, $p = 0.764$, $\eta^2 = 0.007$), probability of win-stay ($F(6.08, 456) = 0.151$, $p = 0.989$, $\eta^2 = 0.002$) and the time to locomote from the nose-poke port to the feeder well ($F(5.59, 419) = 1.39$, $p = 0.220$, $\eta^2 = 0.018$). The response trends are stable within sessions, and are consistent with the univariate analysis on this data collapsed over the session presented above.

3.4. Discussion

Previous rat research has produced inconsistent evidence of sex differences in choice behaviour. In an attempt to clarify this issue, we combined data from 3 past cohorts of animals that followed identical experimental protocols. This allowed us to account for any variance that may have been introduced from different experimenters, testing at different times of the year, or differential auditory environments. Further, this merging of cohorts provided the power to detect slight differences that may have remained insignificant with smaller sample sizes. We analyzed potential sex differences in performance, strategy, and reinforcement sensitivity on a competitive choice task. We found female Long Evans (LE) rats were equivalently likely to engage lose-shift responses as were male LE rats, but were significantly more likely to extraneously sample the reward feeders between trials (EFS). The former is surprising as the literature

suggested females to be more loss-sensitive than their male counterparts (Burton & Fletcher, 2012; Eubig et al., 2014; Jentsch & Taylor, 2003; Lukkes et al., 2016; Orsini et al., 2016; Smethells et al., 2016; van Hasselt et al., 2012). However, as previously discussed, loss-sensitivity refers vaguely to any behaviour directly following the loss event. We have previously shown lose-shift responding is distinct from other mechanisms of reinforcement learning that track reward information over several trials (Gruber & Thapa, 2016), so we cannot rule out sex-based differences on longer time scales. Previous paradigms may have also been more indicative of an emotional reaction to loss or punishment, depending on the task context. Thus, our data suggest male and female LE rats do not differ in their *immediate choice* following reward omission when excluding the EFS effect. The increased EFS in females suggests they may be more susceptible to this potential confound than males in behavioural choice testing.

The females' increased propensity of EFS in the present data may be interpreted as a result of females seeking reward following losses more often than males. However, we found this increase to be independent of whether the animal was rewarded or not, and we did not find females to have an increased probability of lose-shift responding. This leads us to suggest that extraneous feeder sampling is not an immediate result of reward omission in female rats. It is also possible that the increased EFS propensity is more indicative of a lack of effortful control. In humans, however, a substantial meta-analysis examining sex differences in impulsivity (Cross et al., 2011) found no differences in effortful control between men and women. Furthermore, past rodent studies utilizing the delay discounting task, a typical measure of effortful control, have found no baseline sex differences on the task (Eubig et al., 2014; Lukkes et al., 2016; Smethells et al., 2016).

Our lab has also recently altered the length of the barriers separating the reward feeders from the nose-poke port; if EFS was related to motoric effort, we would expect its propensity to decrease with increasing barrier length. However, the rate of EFS increased regardless of an increase or decrease in barrier length (Gruber et al., 2017). Thus, these data suggest this sex difference in EFS is not due to sex-based differences in choice behaviour following reward omission, differences in cost/benefit computations, or differences in motoric effort.

EFS may involve a Pavlovian attraction to the feeder wells. Sex differences have previously been reported in Pavlovian approach (Hammerslag & Gulley, 2014; Pitchers et al., 2015). Our reported increase in EFS by females is also consistent with Pitchers et al.'s finding that female rats made more responses than males in a Pavlovian conditioning task (Pitchers et al., 2015). Many Pavlovian phenomena are modulated by motivation, such as hunger. Interestingly, devaluation of the outcome via pre-feeding before the behaviour did not have an effect on the rate of EFS in male Long-Evans rats (Gruber et al., 2017). These data therefore suggest that non-Pavlovian systems are responsible for EFS in male rats, and this is likely the case for females as well.

We believe that EFS is most indicative of exploration; the sampling of the opposing feeder outside of the task context may be the rodent's attempt to gain more information and explore the environment. Past research utilizing a variety of tasks has consistently demonstrated female rats display more exploratory behaviour than males (Alstott & Timberlake, 2009; Johnston & File, 1991; Lynn & Brown, 2009; Nasello, MacHado, Bastos, & Felicio, 1998; Ray & Hansen, 2004). Many of the tasks in these studies could be confounded by a sexually dimorphic response to stress/anxiety.

However, female rats also show more exploratory behaviour in the novel object recognition task, which has a lower potential for inducing anxiety. (Sutcliffe, Marshall, & Neill, 2007). Furthermore, we have also reported a non-significant trend for females to traverse from the nose-poke port to the feeder well faster than their male counterparts. Although this difference in response time may be due to disparities in body size and weight, it may also be indicative of an increased exploratory drive in the female animals. A sex-based difference in exploration could account for past findings in the IGT. Female rats and humans engage in sub-optimal choice strategies longer than males before ultimately maintaining the optimal choice (Van den Bos et al., 2013, 2012; Van den Bos, Lasthuis, den Heijer, Van der Harst, & Spruijt, 2006). Although this difference is commonly attributed to disparities in loss-sensitivity, it may be more indicative of differential exploratory behaviour in which females may explore more (i.e. require more information) than males to ultimately converge on the optimal choice (Van den Bos et al., 2013). Ethologically, rats face uncertainty in their food source, so there is likely an intrinsic, inextinguishable drive to explore (Gruber et al., 2017). We speculate this drive to reduce uncertainty of food availability may be stronger in females because their smaller weight provides less of a buffer to food deprivation and because fasting negatively impacts reproductive success (Hussain, Tassabehji, Ashton, & Glazier, 2017; Wade, Schneider, & Li, 1996). Indeed, female rats consume more food after fasting than males (Gayle, Desai, Casillas, Beloosesky, & Ross, 2006), suggesting that they are more sensitive to food scarcity. Increased exploration may thus represent a strategy of identifying alternative food sources, to reduce the probability of having no food, in lieu of maximizing caloric intake by focusing responding on a small subset of choice options.

This increased exploration likely has a sexually dimorphic neurobiological cause. We speculate that there are two specific components of the corticostriatal circuit involved: the anterior cingulate cortex (ACC); and dopamine in the nucleus accumbens core (NACc). The ACC is suggested to be involved in the consideration of alternative choices (for review see: (Shenhav et al., 2016)) and plays a significant role in the regulation of exploratory behaviour (Aston-Jones & Cohen, 2005; Quilodran, Rothé, & Procyk, 2008; Weible, Rowland, Pang, & Kentros, 2009). Preliminary evidence from our laboratory indicates that lesions of the ACC reduce EFS. Moreover, there are morphological sex differences in the ACC, particularly in dendritic branching and spine density (Kolb & Cioe, 1996; Kolb & Stewart, 1991; Markham & Juraska, 2002). If these sex-based structural differences influence the function of the ACC or its prevalence in decision-making, it could account for the difference in EFS.

We have previously shown that local microinfusions of d-amphetamine into the NACc increase EFS in male LE rats (Wong, Thapa, et al., 2017), suggesting that increased dopamine (or other amphetamine-affected catecholamine) in this region promotes EFS. Female Long Evans rats are also more sensitive to amphetamine than their male counterparts, exacerbating drug-induced changes in decision-making tasks (Eubig et al., 2014; Orsini et al., 2016). Females have increased dopamine receptor levels and availability of dopamine throughout the striatum (Becker et al., 2012; Mozley, Gur, Mozley, & Gur, 2001; Walker, Ray, & Kuhn, 2006; Walker, Rooney, Wightman, & Kuhn, 2000). Linking these observations thus provides one possible explanation for heightened EFS in females as compared to male rats. This increase in dopamine transmission is particularly pronounced when under the effects of increased estrogen, as

is experienced during proestrous (Datla, Murray, Pillai, Gillies, & Dexter, 2003; Lammers et al., 1999; Pasqualini, Olivier, Guibert, Frain, & Leviel, 2002; Torres-Hernández & González-Vegas, 2005; Zhang, Yang, Yang, Jin, & Zhen, 2008). Estrous cycle differences have also been shown to affect exploratory behaviour and alter task strategy (Korol, Malin, Borden, Busby, & Couper-Leo, 2004; Tropp & Markus, 2001). We did not explicitly test or control for estrous cycle, but we expect that the female rats used in our study were randomly cycling during testing. Our positive finding of sex-based difference in this response element indicates that the sex difference is sufficiently robust and/or large to not depend on rigorous control of this independent factor. Furthermore, previous rat research has shown no alterations in performance on the rodent version of the Iowa gambling task, or rodent tests of impulsivity based on estrous-related hormonal fluctuations (Jentsch & Taylor, 2003; Lukkes et al., 2016; Peak et al., 2015; Smethells et al., 2016; Van den Bos et al., 2013, 2012). In sum, we speculate that increased dopamine transmission in the NACc and/or increased utilization of the ACC in decision-making by females, as compared to males, could account for their increased engagement of exploratory EFS behaviour.

3.5. Conclusions

We found no sex-based differences in the immediate (trial-by-trial) adaptation of choice following wins or losses when controlling for confounding factors. This argues against the notion of a sex difference in general 'loss-sensitivity.' We did find a robust sex-based difference in feeder sampling between trials, which could reflect exploration. This between-trial behaviour affected subsequent choice, and thus presents a confound in the study of choice, particularly because it was more prevalent in female rats. After

controlling for this confound, we found no sex difference in the number of rewarded trials, win-stay responding, or lose-shift responding, suggesting similar decision-making *performance*. Thus, our study highlights the need for future paradigms to be cognizant of differential exploratory behaviour, both within and between trials, and its influence on subsequent choice.

4. Distributed encoding of reinforcement in rat cortico-striatal-limbic networks

4.1. Introduction

The control of voluntary actions in the mammalian brain appears to be distributed among multiple neural circuits involving cortical, striatal, thalamic, and limbic structures (Burton et al., 2015; Floresco et al., 2008; Gruber & McDonald, 2012; Haber, 2003). The interaction of these circuits facilitates learning in many contexts, from ethologically primitive grooming behaviours, to highly complex and abstract behaviours, such as choosing a spouse, or making a political choice (Berridge & Whishaw, 1992; Mendez, 2017; Ruff & Fehr, 2014). Despite some segregation of function among the structures, neurons responding to reinforcement outcomes may be found throughout the interconnected network of these structures. In rodents, the fraction of neurons encoding reinforcement ranges from approximately 15% in the ventral striatum (VST) and basolateral amygdala (BLA), to 25-45% in the medial prefrontal cortex (mPFC), 37% in the dorsomedial striatum (DMS), 45% in the dorsolateral striatum (DLS), and 53% in the orbitofrontal cortex (OFC; Pratt and Mizumori, 2001; Roesch et al., 2006, 2009, 2010; Sul et al., 2010; Kim et al., 2013; Atallah et al., 2014; Shin et al., 2018). These proportions are similar to those reported for homologous structures in primates (Kobayashi et al., 2006; Lau & Glimcher, 2007; Simmons, Ravel, Shidara, & Richmond, 2007; Sugase-Miyamoto & Richmond, 2005). However, a larger proportion (26%) of neurons respond to reward in the primate ventral striatum (Simmons et al., 2007). The presence of this reward signal throughout the network may be due to the interconnectivity among the structures, and is possibly facilitated by the widely distributed dopaminergic projections from the midbrain. These dopaminergic neurons provide information related

to reward prediction errors (Bayer & Glimcher, 2005; Hollerman & Schultz, 1998; Schultz, 2016), motivational salience (Berridge & Robinson, 1998; Bromberg-Martin et al., 2010), and locomotion initiation (Howe & Dombeck, 2016).

These observations of widespread reward signalling pose several important questions: does processing of reward information begin in one brain region and then project outward, or does the signal encoding reward emerge in multiple regions at once? Further, do any individual brain structures contain reward-related information that cannot be garnered from any of the others? In attempt to generate data bearing on these questions, we recorded field potentials (FPs) from seven distinct regions of the cortico-striatal-limbic network in rats performing a binary choice task with intermittent reward delivery (Donovan et al., 2018). We utilized methods from machine learning and signal processing to quantify the relative information encoded by brain structures individually or in sets. This methodology allowed us to explore the encoding of the reinforcement signal from intact animals, without making assumptions about prevailing frequency ranges or specific contributions by individual brain regions. Within the FPs of each region, we found a high degree of overlap in reinforcement information, suggesting widespread encoding of this fundamental aspect of reward processing. This analysis reveals several key parallels with that of human electroencephalogram research, providing evidence that immediate processing of reinforcement information may be relatively conserved across species.

4.2. Methods

4.2.1. Animals and Experimental Design

The dataset included 5 adult male Long Evans rats (515 – 660g). These data were previously collected as part of a larger pharmacological study (unpublished) and we have taken data from the control group of this past study for the present analysis. Consequently, each rat received an injection of 0.1mL saline on alternating days over 2.5 weeks, for a total of 11 injections. These injections, however, occurred during the initial pre-training on the behavioural task, and injections ceased 9 days prior to the data used in the present analysis.

Prior to surgery, rats were initially shaped with every feeder well choice rewarded for 5 – 8 days until they reached a criterion of at least 100 trials over a 45-min session. Following surgical recovery, rats were retrained with every response rewarded until they met a criterion of 150 trials per 45-min session. Rats completed two more 45min training sessions with rewards present in both feeder wells. In subsequent sessions, rats completed 45min testing sessions in which an algorithm attempted to minimize the number of rewards given. This algorithm predicted, based on the choices of the previous 4 trials, which feeder the rat would select; choice of the opposing feeder was then rewarded. Thus, optimal performance was random feeder well selection, which would result in reward on 50% of trials. Rats completed a total of 5 testing sessions on this competitive protocol. Field potential recordings were obtained from 3 of these sessions. Thus, a total of 15 sessions were recorded. Due to signal complications resultant from recording in freely moving rats, data from 5 of these sessions, from 4 different animals, were used for the analysis detailed here.

4.2.2 Surgeries

Each rat first received an injection of the opioid analgesic buprenorphine (0.03mg/kg s.c.). Thirty minutes later, anesthesia was induced using isoflurane (5% for induction and 1.5–3% for maintenance) and the animal was placed in a stereotaxic frame. Small craniotomies were made in the skull for electrode implantation. Teflon-coated tungsten electrodes (76-µm diameter; A-M Systems) were implanted unilaterally (right side) in the following brain regions relative to bregma (Fig. 5; in mm; AP, ML, DV): mPFC (3.2, 0.6, 2.6), BLA (-2.3, 5.0, 8.4), VHC (-4.9, 5.1, 7.2), VST (2.0, 1.6, 64), DMS (0.84, 1.9, 3.5), OFC (4.2, 2.2, 4.0) and DLS (0.84, 3.6, 4.5).

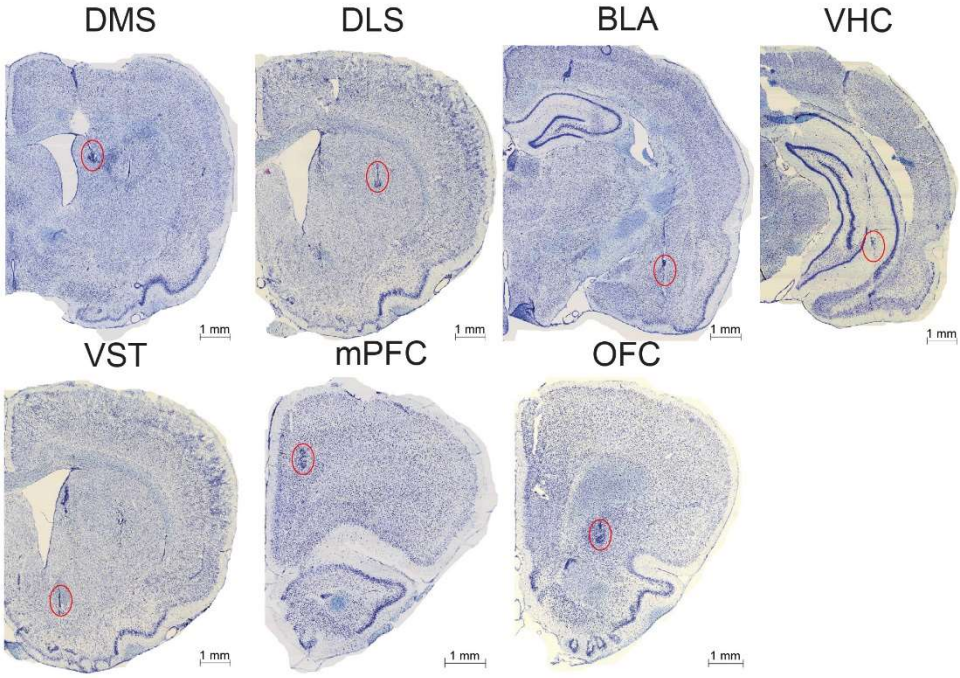

Figure 5. Representative electrode placement. Coronal sections of Nissl-stained rat brains. Current was passed through the recording electrode to induce a lesion (circled in red) marking the electrode locations. Abbreviations: dorsomedial striatum (DMS), dorsolateral striatum (DLS), basolateral amygdala (BLA), ventral striatum (VST), ventral hippocampus (VHC), medial prefrontal cortex (mPFC), orbitofrontal cortex (OFC).

Ground and reference screws were inserted in the skull over the cerebellum and four anchor screws were implanted laterally. Dental acrylic was applied to secure an

electrical connector (A-M Systems #520100 & 520200; Ginder Scientific GS09PLG-220 & GS09SKT-220) to the skull in order to interface implanted electrodes with a head-mounted preamplifier during experiments. Postoperatively, rats were treated with the corticosteroid analgesic Metacam (1 mg/kg s.c. for 3 days) and antibiotic Baytril (10 mg/kg s.c. for 5 days). The rats were allowed to recover for at least 7 days before retraining on the behavioural task.

4.2.3 Field Potential Recording

FP recording followed the procedure of a previous study (Skelin, Needham, Molina, Metz, & Gruber, 2015). A tether cable, connected to a pre-amplifier 'headstage' (Neurlaynx, HS-27M), was secured to the head-mounted connector when animals were placed in the behaviour box. Field potential signals were recorded using a Cheetah acquisition system (Neuralynx, Bozeman, MT, USA). The signals were filtered at 0.1–600Hz and digitized at a 2kHz sampling rate. The signal from each channel was referenced to the ground screw. After the conclusion of behavioural testing, anodal current (10μA for 30s) was passed through each of the recording electrodes to induce marking lesions around the electrode tips. Seven days later, the rats were transcardially perfused with phosphate-buffered saline, followed by 4% paraformaldehyde. After fixation in 4% paraformaldehyde for 24h and dehydration in 30% sucrose for three days, the brains were coronally sectioned with a cryostat (40-micron thickness). Sections were then mounted, dried and labelled for Nissl bodies. Placement of electrodes was identified using light microscopy and referenced against an anatomical atlas (Paxinos & Watson, 2005).

4.2.4. Signal Preprocessing

Signal preprocessing was completed using MATLAB 2018b (The MathWorks Inc., 2018). Recorded signals were parsed into individual events by creating an analysis window beginning 2s prior to feeder entry, and ending 3s after this event. Feeder entry was defined as the time of infrared beam break, and this served as a proxy for the time of reward delivery, or omission. This signal was first de-trended using the function "locdetrend" from the Chronux toolbox (Mitra, 2007) using a 0.05s window with 0.01s overlap. To remove low frequency fluctuations of the FP related to movement, the signal was band-pass filtered from 3 – 110Hz using an FIR filter (designed using the MATLAB function firls) applied using the MATLAB function filtfilt. To remove line noise, a FIR band-stop filter from 56 – 64Hz was applied using filtfilt. Electrical artifacts, defined as those points greater than five standard deviations from the mean signal, were removed and the missing values were interpolated using the inpaint_nans function (D'Errico, 2006). The signal was then band-pass filtered again using an FIR filter from 3 – 110Hz with filtfilt to remove any artifacts added by interpolation. This resulted in the interpolation of signals in 0.69% of reward omission recordings, and 2.4% of rewarded recordings. Following this processing, any trials that still contained a point greater than five standard deviations from the mean signal or contained a clipped signal for greater than 50ms were removed. This preprocessing resulted in a final sample size of 857 trials, consisting of 423 rewarded trials and 434 unrewarded trials.

4.2.5. Neural network analysis and statistics

Using the Python programming language, signals were down-sampled by a factor of 4 to a sampling rate of 500Hz. The z-score was then taken of each 5s segment for each area and a spectrogram was generated for each z-scored signal using the Python Matplotlib function specgram at 250 time steps per segment, with a 248 time step overlap (Caswell et al., 2018). The time window from 0 – 0.5s was taken from this spectrogram for further analysis; 0s represented the time at which the rat broke the infrared beam that recorded feeder well entry. Thus, this window is the 500ms immediately following reward delivery or reward omission. To ensure the classifier was trained using a chance accuracy of 50%, a small number of trials from the lacking condition, chosen randomly, were repeated in the dataset; this resulted in the repetition of 11 rewarded trials to equal the number of non-rewarded trials.

An artificial neural network classifier was constructed using Keras (Chollet, 2015) with a Tensorflow backend (Abadi et al., 2016). This network consisted of a 2-D convolutional layer, consisting of 32 filters, with a kernel size of 2x3 and a stride size of 1x2, followed by a batch normalization layer and a rectified linear unit activation function. In order to overcome the degradation of training accuracy with increasing layer count, the convolutional layer was followed by 2 residual unit blocks (He, Zhang, Ren, & Sun, 2016). Each block consisted of 2 2-D convolutional layers, with the input to each block added to its output. Each block was then followed by batch normalization and a ReLU activation function. The first residual block's convolutional layers used 16 filters, the first with a kernel size of 1x1 and stride size of 1x1; the second with a kernel size of 3x3 and stride size of 3x1. This block was followed by a 1x2 max pooling layer. The

second residual block's convolutional layers used 16 filters, the first with a kernel size of 1x1 and stride size of 1x1; the second with a kernel size of 3x3 and stride size of 1x1. This block was followed by a 2x1 max pooling layer. Dropout was then applied at a probability of 0.55. The output layer then consisted of a single fully connected neuron with a sigmoid activation function. This model was then fit to the data over 80 epochs using a binary cross entropy loss function and the Adam optimizer with a batch size of 16. The learning rate was decremented by a factor of 10, to a minimum of 10^{-8} if the test set loss did not decrease after 6 consecutive epochs. This was repeated 10 times, each using 90% of the data as a training set, and 10% as a test set. The average of these 10 folds is reported as the final classification accuracy.

4.2.6. Post-hoc Analysis

One-way ANOVA tests and two-sample t-tests were implemented using MATLAB (The MathWorks Inc., 2018). Any groups that failed the assumption of variance used the Satterthwaite approximation for degrees of freedom (Hall & Willink, 2001).

To visualize the activations of the neural network, we used the "grad-cam" function of the keras-vis python toolbox (Kotikalapudi, 2017). This function generates a class activation map for the output of the final convolutional layer by backpropagation through the network for each test set example (Selvaraju et al., 2017). We took the mean of all test set examples for all ten folds to generate the class activation maps that are shown.

To calculate Granger causality, we utilized the multivariate Granger causality toolbox (Barnett & Seth, 2014). As this analysis may be distorted by over processing, the only processing performed was removal of recordings with no signal for greater than 200 time steps (100ms). This change resulted in a total of 1154 trials from 4 animals from a total of 5 sessions. To fit the vector autoregressive model, the best model order fit was calculated using Akaike information criterion to be 15. We computed both the Granger causality in the time domain and frequency domain via this toolbox.

4.3. Results

4.3.1 Reinforcement information is encoded by all region-specific FPs

We simultaneously recorded field potentials from 7 brain structures (Fig. 6A) in freely-moving rats performing a binary choice task. The optimal strategy on this task was for the rat to maintain an unpredictable choice selection from trial to trial. Across all sessions, rats received reward on 46.7% ± 3.5% of trials, which is near the maximum expected value of 50%. Beginning at the time of reinforcement, a 500ms time window was chosen for analysis. The beginning of the window corresponded with the opening of the liquid sucrose feeder valve in rewarded trials, or the illumination of a house light in unrewarded trials. This time window duration was chosen in an attempt to avoid potential effects of differential licking on the FP among reinforcement types. Visual inspection of the lick histogram for rewarded trials, contrasted with unrewarded trials, showed that licking drastically diverged between outcomes 0.5s after the reinforcement onset (Fig. 6B). Unless otherwise stated, subsequent analysis was restricted to the period from 0 to 0.5 seconds after reinforcement onset. The time-frequency spectrograms of the FP during this time window were calculated for each brain area (Fig. 6C) and combined to form a

matrix of size 61 x 124 x 7 (frequency bins x time bins x brain regions). The matrices were given as input to a neural network classifier, with the goal of classifying a trial as a "win" (the rat received reward at its chosen feeder well), or a "loss" (the rat did not receive reward at its chosen feeder well). Each reported accuracy value is the result of 10-fold cross validation.

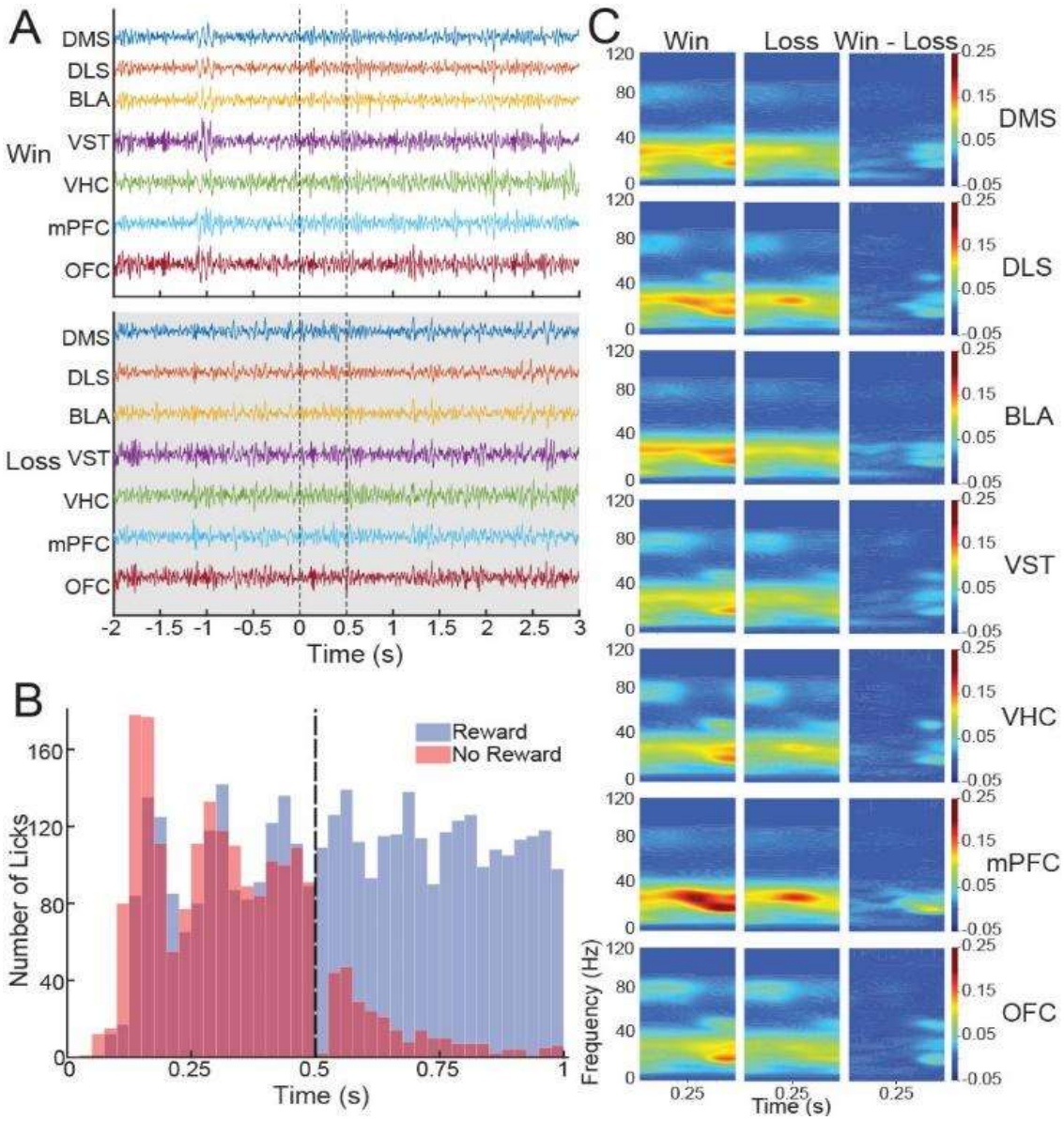

Figure 6. Electrophysiological signals and licking behaviour. A) Simultaneous field potential recordings from 7 brain areas during a randomly chosen rewarded (top) and unrewarded (bottom) trial. Traces are aligned to reward (t=0), and the time window used for analysis is marked by the vertical dashed lines. B) Histogram of licking events in feeders in rewarded and unrewarded trials, aligned to reinforcement. C) Mean FP spectrograms for each brain region in rewarded trials (left), unrewarded trials (middle), and their difference (right). Abbreviations: dorsomedial striatum (DMS), dorsolateral striatum (DLS), basolateral amygdala (BLA), ventral striatum (VST), ventral hippocampus (VHC), medial prefrontal cortex (mPFC), orbitofrontal cortex (OFC).

When given the data from all seven recorded brain areas, the network predicted trial outcome with an accuracy of 76.9% ± 3.4%. We refer to this value as the baseline because it uses all available information from all of the recording sites. We then

investigated which brain areas' FP contained information useful for this prediction. To test this question, we trained and tested the classifier using all but one of the recorded brain regions. There was no significant difference (at $\alpha = 0.05$) between prediction accuracy achieved when data from any one structure were excluded, and the accuracy achieved at baseline (when all data were included). There was also no significant difference between accuracies computed for the groups of six FPs, no matter which area was excluded (Fig. 7A; ANOVA: $F_{6, 63} = 0.44$, $p = 0.85$). This result suggests no single recorded region encoded reinforcement information that could not be garnered from the collective data of the remaining regions. We then conducted the inverse analysis, excluding data from all but one area. In this case, accuracy was significantly lower than baseline (mean difference = -11.3% ± 2.3%, mean Bonferroni $p = 0.001$), suggesting a single brain region's FP within this time window does not include all of the relevant reward information of this neural circuit. There was a significant difference between group means when single-region data were used for classification (Fig. 7B; ANOVA: $F_{6, 63} = 2.3$, $p = 0.047$). Post-hoc analysis revealed this difference was due to one pair; the classification accuracy that used data from only the OFC was significantly lower than the accuracy when using only the data from the DMS (Fig. 7B; Tukey-Kramer: $p = 0.024$). These results suggest that no single area contains all of the information available in the network.

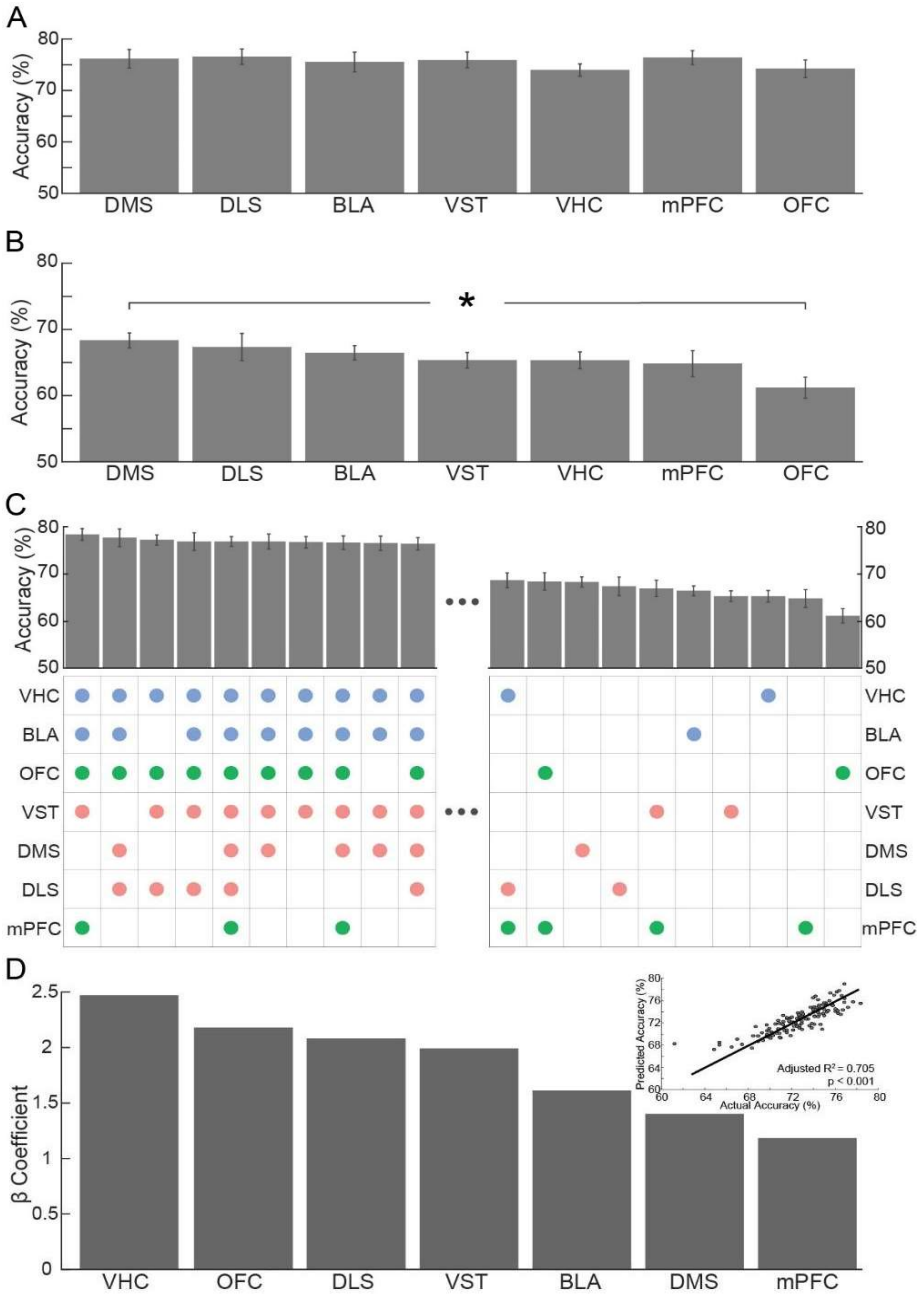

Figure 7. Classification of reinforcement from FP spectrograms. A) Classifier accuracies (10-fold cross validation) for FP from all but one of the recorded regions, showing that no single brain regions was required for maximal accuracy. The axis labels indicate which brain region was excluded from the data. B) Classifier accuracy from individual brain areas, revealing that each area is less predictive than the full ensemble of structures. C) Classification from all combinations of structures. The top 10 (left) and bottom 10 (right) 10-fold cross validation accuracies are shown. Coloured dots below each bar indicate which regions' field potentials were included. D) Regression of accuracy using brain area inclusion as predictors. The β coefficients indicate the relative contribution of each region's FP to the accuracy of classification both individually, and in combination with field potentials from other regions. The predicted accuracies are plotted against actual accuracies (inset) as a validation of the regression model. Error bars indicate standard error of the mean.

Next, we tested information encoded by sets of region-specific FPs by training and testing the classifier on every combination of the FP data from the seven recorded areas. This resulted in a total of 127 different combinations of data. For brevity, we have shown the top ten and bottom ten, sorted by descending accuracy (Fig. 7C). FPs from the VHC were the most prominently featured in the top ten combinations with highest prediction accuracy. The BLA, OFC and VST were each included in all but one of the top ten accuracy results. Conversely, data from the mPFC were excluded most often in the top ten accuracies, and included most often in the bottom ten. These data suggest the VHC, BLA, OFC and VST encoded the most relevant information for classification of reward in this choice task, while the mPFC encoded the least. However, the brain regions involved in these top 10 and bottom 10 classification accuracies were not concordant with accuracies achieved when using each brain regions' FP individually. Although these disparities in accuracy are small, they suggest the VHC, BLA, OFC, and VST provide information that is a weak predictor on its own, but synergizes information from other regions to produce higher levels of classification accuracy.

To further investigate the classification accuracy disparity between FPs from individual areas from those from sets of brain areas, we performed a linear regression employing all brain region combinations (adjusted $R^2 = 0.71$, $F(7, 119) = 44.0$, $p < 0.001$). Specifically, we formed a 127 element vector for each combination, where areas included in a combination were assigned a value of 1 and areas not included in a combination were assigned a value of zero. This regression therefore reveals the contribution of each brain region to the achieved classification accuracies among all combinations. Although we recognize that there are likely unavoidable issues with

correlations among predictive factors, this analysis provides an objective measure of the information gained from each region (Fig. 7D). Consistent with the accuracy achieved when data from the mPFC were used individually, and its prevalence in the bottom 10 combinations, data from the mPFC resulted in the smallest regression coefficient, but were nonetheless a significant predictor of variance ($\beta = 1.40$, $t = 5.0$, $p < 0.001$). Surprisingly, the VHC data's regression coefficient was the largest of all the recorded regions ($\beta = 2.5$, $t = 8.9$, $p < 0.001$). This result is in concordance with the prevalence of the VHC in the top 10 combinations and suggests the VHC may help bias the classifier when its data is used in combination with other regions. Data from the OFC resulted in the second largest regression coefficient ($\beta = 2.2$, $t = 7.8$, $p < 0.001$), despite achieving the lowest accuracy when used individually for prediction. The regression coefficients of the DLS ($\beta = 2.1$, $t = 7.5$, $p < 0.001$) and VST ($\beta = 2.0$, $t = 7.1$, $p < 0.001$) were similar to the OFC, suggesting the FPs from these 3 areas may also encode information complementary to the other recorded regions. The BLA's regression coefficient was intermediary to the other regions ($\beta = 1.6$, $t = 5.8$, $p < 0.001$), which is in agreement with the classifier's achieved accuracy when this area's data were used individually for prediction. Conversely, FPs from the DMS resulted in the second lowest regression coefficient ($\beta = 1.4$, $t = 5.0$, $p < 0.001$), suggesting the information encoded by the DMS, although individually the most predictive of reinforcement, provides little additional information.

4.3.2 Time-Frequency features informing reinforcement

Next, we sought to identify what features of the FP spectrograms were important for classification. An activation map (Zhou, Khosla, Lapedriza, Oliva, & Torralba, 2016)

was used to visualize which regions of the spectrograms were preferentially utilized by the classifier. A map was computed for each trial in the test set and for each fold of the 10-fold cross validation, then averaged (Fig. 8A). Higher activation scores indicate the feature of the input had a stronger impact on performance than other features. When the data from all areas were included, the classifier's mean activations were predominantly in the 5 – 35Hz range, with the largest magnitude in the last 100ms of the time window (Fig. 8A, 1st panel). When using data from only one area for classification, the classifier's activations were similar, with the largest activation between approximately the 5 – 35Hz range (Fig. 8A, 2nd - 7th panel). However, there were differences in the dominant time periods. Specifically, activations from the VHC, DLS, DMS and VST were predominantly in the last 200ms of the time window, whereas activations from the OFC and mPFC were more diffuse throughout the window. Classifier activations due to data from the BLA predominated in the first 200ms. To further investigate these results, we masked (i.e. set to zero) time-frequency windows in the source spectrograms prior to training & testing the classifier (Fig. 8B). Masking the second half of the time window caused a significant reduction in classification accuracy below baseline (Tukey-Kramer post-hoc test: $p = 0.001$). Masking the 0 – 40Hz range across the entire time window did not result in a significant decrease in accuracy (mean accuracy decrease: 3.86%, Tukey-Kramer post-hoc test: $p = 0.98$). However, masking the 40 – 120Hz range resulted in a decrease in accuracy near statistical significance (mean accuracy decrease: 7.5%, Tukey-Kramer post-hoc test: $p = 0.051$). In sum, although the classifier was maximally activated by the frequency ranges below 5 - 35Hz, this range was not necessary to predict task outcome, whereas the 40 – 120Hz range was important for classification.

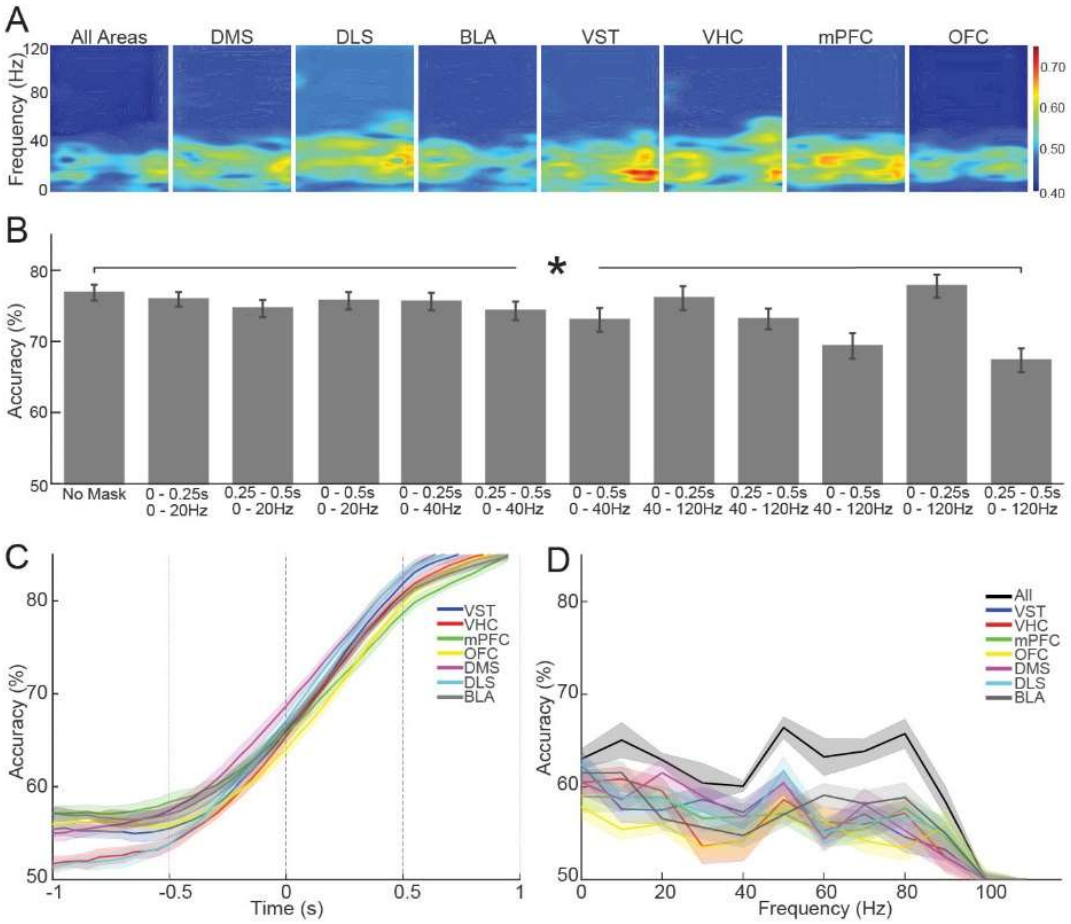

Figure 8. Time-frequency features of FP predictive of reinforcement. A) Mean activation maps of the classifier when trained on all areas (far left panel), or individual brain regions (as marked). The amplitude (colour) represents a normalized value of the classifier's internal activations, and reflects the salience of each time-frequency bin. B) Classification accuracy when data from all areas are presented with time-frequency ranges masked (as indicated in axis label). Masking all frequencies in the second half of the time window of interest significantly reduced accuracy below baseline, and masking frequencies above 40Hz across the entire time window of interest reduced accuracy below baseline to near statistical significance ($p = 0.051$). C) Classification accuracies (moving average) based on data selected by a sliding time window (0.5s) starting 1 second prior to reinforcement (t=0). Each point represents the achieved classification accuracy using data within a 0.5s window following that time point. D) Classification accuracies of a 0-0.5s time window and sliding bin of frequency (10Hz). Error bars indicate standard error of the mean.

We next tested if a specific brain area may lead the reward encoding. We repeated the classification analysis on a sliding window of data (500ms duration, 100ms increment) starting from one second *prior* to the time of reward (Fig. 8C). The resulting spectrograms for each brain area were used as the input to the classifier. The first window generated information from 1 - 0.49s prior to reinforcement. In the -0.5 to 0s time range,

FPs from the DMS appeared to achieve slightly higher classification accuracies than the others. This result suggests the DMS may encode reinforcement prior to the other areas of interest. Although the network was given an equivalent amount of rewarded and non-rewarded trials, we found the classifier's accuracy remained above that of chance (50%) prior to any window that included the reinforcement event. We suspect this resulted from information about the prior probability of reward being encoded by FPs. Animals were rewarded on 46.7% of trials, so simply guessing that no reward would arrive would be correct on 53.3% of trials. Although we duplicated trials from the underrepresented outcome (11 trials) to ensure an equivalent number of rewarded and unrewarded trials in the dataset, the FP prior to the reinforcement event may have encoded the probability of reward received in that session. However, this probability tracking may not have been present in all regions. FPs from the VHC and DLS resulted in chance accuracy when using any time window prior to the inclusion of the reinforcement event for classification. This result suggests the BLA, DMS, VST, mPFC and OFC all have access to some expectation of reward.

The above analysis revealed the 40-120Hz frequency range was necessary for good classification performance. We sought to determine if a narrower frequency band was responsible for this finding. We therefore repeated the classification procedure using a sliding frequency band of 10Hz from the FPs of each individual recorded area, and the FPs of all areas together, by masking the remaining frequencies in each spectrogram (Fig. 8D). Using FPs from all the recorded regions, classification accuracy was significantly lower than baseline for every 10Hz band tested (Bonferonni-Holm corrected p-values < 0.01). Although still significantly lower than baseline (Bonferonni-Holm corrected p-

value: 0.003), the 50-60Hz frequency band achieved the highest classification accuracy of all the 10Hz bands. A similar trend was observed in the accuracies achieved from these 10Hz bands with FPs from individual areas. These results suggest the necessity of frequencies above 40Hz for good classification accuracy is not due to a narrow band (<10Hz), but rather a broader frequency range.

4.3.3 Information flow among structures

The previous analysis provided little evidence for temporal order of reinforcement encoding among the regions of interest. Specifically, only the DMS appeared to have slightly higher accuracy than other regions in the -0.25 to +0.25 s window (Fig. 8C). We therefore utilized a method (Granger causality) specifically designed to uncover temporal relationships with high temporal precision (Fig. 9). This method quantifies the effective connectivity of one signal source on another, which provides a measure of signal transmission directionality in the network (Bressler & Seth, 2011). Due to the large number of trials we utilized in this analysis, all of the Granger-causality values were significant at $\alpha = 0.05$ (Granger's F-test). The strongest interactions were VHC -> DLS and OFC -> VST, followed by the inverse of these pairs (Fig. 9A). A Granger causality analysis was performed in the frequency domain to identify which frequency bands were important for signal transmission (Fig. 9B). The strongest Granger causalities were observed in the DLS -> VHC (below 60Hz) / VHC -> DLS (above 60Hz) and the VST-> OFC (below 60Hz) / OFC -> VST (above 60Hz). The inverse nature of these pairings suggests a reciprocal relationship between these structures. Further, FPs recorded at the DMS Granger-caused nearly all the other areas of interest below approximately 30Hz.

This widespread directionality suggests a functional connectivity pattern in which the DMS serves as a central hub.

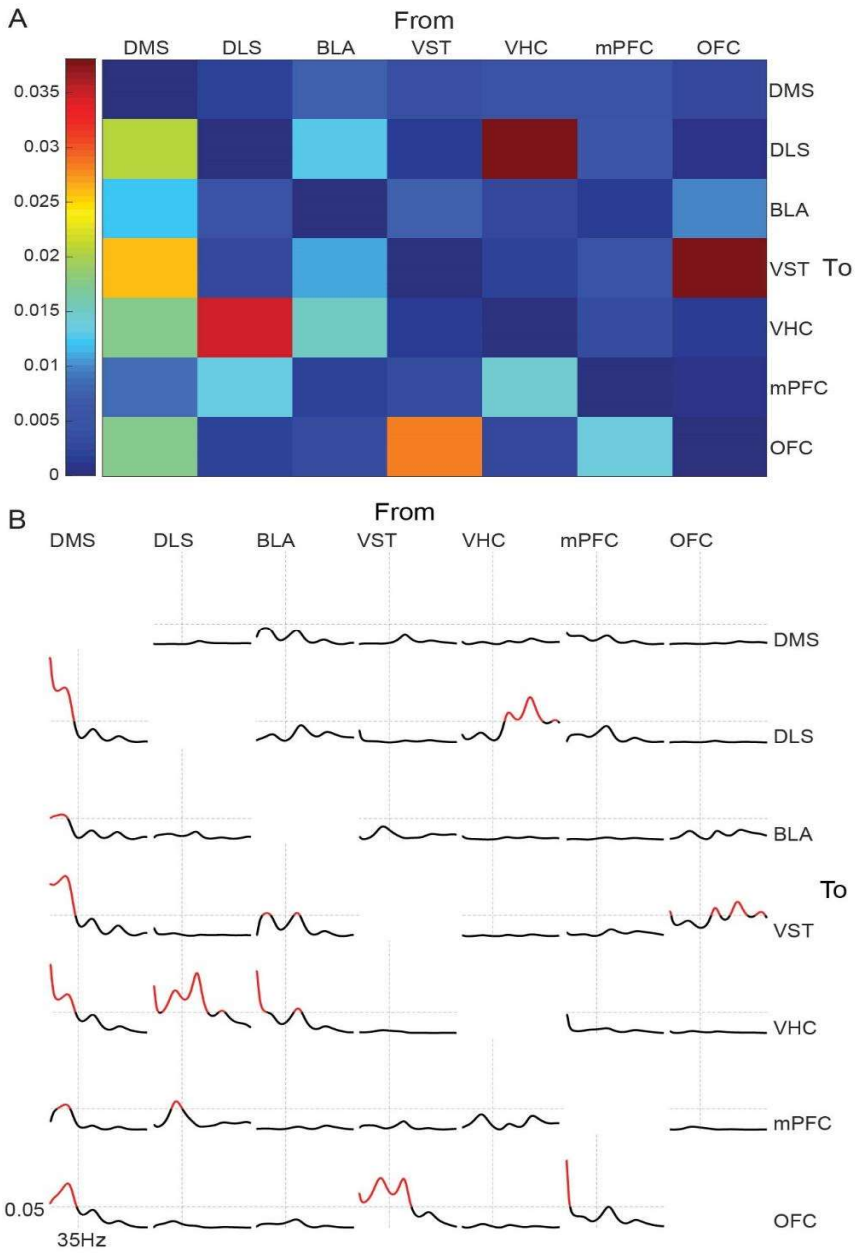

Figure 9. Analysis of directed information transfer. A) Granger causality among brain regions, revealing information transfer between: DMS -> DLS, VST, VHC and OFC; DLS -> VHC; VHC -> DLS; OFC -> VST and VST -> OFC. All granger-causality results shown are statistically significant at $\alpha = 0.05$. B) Granger causality decomposed by frequency. Values above the significance threshold (0.05) are highlighted in red. The DMS shows significant information transfer to nearly all the other recorded regions at frequencies below 35Hz. The DLS shows granger causality to the VHC below 60Hz, and the VHC shows granger causality to the DLS above 60Hz. The VST shows granger causality to the OFC below 60Hz and the OFC shows granger causality to the VST above 60Hz.

4.4. Discussion

Loss-of-function studies have enabled researchers to suggest many task-specific roles to individual regions of the decision-making system (for review see: Doya, 2008). However, these experiments often assume functional modularity, and cannot account for possible compensation by non-target regions (Rorden & Karnath, 2004). Although we recognize methodological problems inherent to interpreting field potentials, such as source attribution (Herreras, 2016), the use of FPs permits investigation into intact functioning of this circuit. The collective FP from the recorded areas immediately after reinforcement was predictive of outcome on 76.9% of trials. No individual brain region was necessary to achieve this accuracy, suggesting information related to outcome valence is distributed throughout the afferents of all the recorded regions. Further, our results demonstrated FPs from any of the individual recorded areas could predict reinforcement outcome. This pervasiveness may account for the inability of past research to ascribe more general features of reward processing, such as reinforcement valence, to any individual region or combination of regions. Vickery et al.'s (2011) research suggested this ubiquitous reinforcement processing in humans using fMRI, which is qualitatively similar to the ubiquity in rats we describe here. Although ubiquity in reinforcement outcome encoding has been shown in the rodent cortex (Sul et al., 2010), we believe we are the first to demonstrate widespread reinforcement encoding throughout the cortico-striatal-limbic system in rodents.

FPs are suggested to reflect the activity of afferents to the respective regions, rather than activity of neurons whose somata reside in the named locale (Herreras, 2016). For instance, the DMS was Granger-causal of all the other recorded regions below ~35Hz

in the present data. Past research has proposed the DMS as an integration point for multimodal sensory input (Reig & Silberberg, 2014). This structure also receives input from the mPFC, OFC, BLA and intrastriatal connections (Schilman, Uylings, Graaf, Joel, & Groenewegen, 2008; Voorn et al., 2004). Furthermore, the DMS, and all of our regions of interest, receive projections from dopamine neurons in the midbrain (Björklund & Dunnett, 2007; Voorn et al., 2004), which have been shown to modulate their activity based on reinforcement outcome (Schultz, 2016). The FPs we recorded from the DMS are likely indicative of some combination of incoming information to the structure, including afferents from the PFC, VHC, BLA, and midbrain dopamine neurons. Thus, the widespread Granger-causality of the DMS reported here may be resultant of the influence of these inputs to DMS, rather than vice versa.

We found the data recorded from the DMS to be most predictive of trial outcome, whereas recordings from the OFC were least predictive. This discrepancy is consistent with past literature suggesting the OFC encodes *expected* outcomes, and neurons predominantly fire preceding the reinforcement (Schoenbaum et al., 2009). The actual reward in our task is primarily stochastic, and so the anticipation should be largely uncorrelated with actual outcome, except for a small amount of prior information, given that animals received reward slightly less than chance (46% vs 50%). Furthermore, the DMS has been shown to encode an action-specific reward prediction error, which has not been observed in the ventral or dorsolateral subdivisions of the structure (Roesch et al., 2009; Stalnaker et al., 2012). However, when FPs from multiple regions were given as input to the classifier in combination, FPs from the DMS in these combinations were nearly the least predictive of reinforcement. This disparity between individual and

combinative contributions of the DMS FPs to trial classification suggests this region, in relation to the other recorded regions, encodes relatively redundant reinforcement information.

Previous researchers have reported that hippocampal neurons responded to reward, as a consequence of run speed differences as the animal reached the reward site (van der Meer, Johnson, Schmitzer-Torbert, & Redish, 2010). Our data suggest a counterexample, because the animals are mostly stationary at the reward feeders. We found FPs from the VHC were most predictive of reward if utilized in combination with data from other regions. The VHC likely provides complementary information to the other regions, which influences their encoding of reinforcement. This finding agrees with past suggestions that the VHC provides non-spatial, contextual information to the mPFC and striatum (Euston et al., 2012; Pennartz et al., 2011). Similarly, we found FPs from the OFC and VST to be more predictive of reward in combination with data from other regions, than when analyzed individually. The similarity found between these areas is not surprising. These regions both receive extensive innervation from reward-related, midbrain dopaminergic neurons (Burton et al., 2015; Murphy & Deutch, 2018). Thus, information encoded in the afferents of one of these areas is likely to be found in the other. We also found a reciprocal Granger-causality between the OFC and VST, with the VST Granger-causing the OFC below ~50Hz, and the OFC Granger-causing the VST above ~50Hz. This frequency specific reciprocity is consistent with previous data in which VST stimulation has been shown to enhance low frequency OFC FPs (McCracken & Grace, 2007) and entrainment of OFC neurons to gamma oscillations has been shown during odor sampling (Wingerden, Vinck, Lankelma, & Pennartz, 2010). Interestingly,

FPs from the mPFC in combination with those from other regions achieved the lowest accuracies in predicting reinforcement. Previous research suggests the mPFC is involved in contextual associations (Barker et al., 2007), temporal representation of reward contingency (Coutureau et al., 2012), comparison of cost and expected value of choices (Shenhav et al., 2016), and assessment of counterfactual outcomes (Mashhoori et al., 2018). Thus, although the mPFC may perform subjectively higher order functioning in relation to reinforcement, this information may not aid in encoding the valence of delivered reinforcement. Overall, the disparities in predictive power when brain regions' FPs were analyzed individually, rather than in combination, suggest the presence of both parallel and integrative processing in this circuit. This proposal has previously been made in Haber's (2003) seminal review on the primate basal ganglia; yet, we are surprised to find evidence for this in such a fundamental aspect of reward processing as reinforcement valence.

Activation maps of the model's output layer showed the 5 – 35Hz range as the most prevalent for outcome classification across all recorded regions. This frequency range coincides with human EEG studies in choice tasks. Differences in theta (4-8Hz) and beta-gamma (12-35Hz) power occur at the time of outcome (Cavanagh, Figueroa, Cohen, & Frank, 2012; HajiHosseini & Holroyd, 2015; HajiHosseini, Rodríguez-Fornells, & Marco-Pallarés, 2012; Marco-Pallares et al., 2008; Marco-Pallarés, Münte, & Rodríguez-Fornells, 2015; Mas-Herrero & Marco-Pallarés, 2014). However, our masking of frequencies below 40Hz did not significantly reduce accuracies below baseline, whereas masking of frequencies above 40Hz resulted in reduced accuracy that neared statistical significance ($p = 0.051$). Increased power within the lower frequency range has

previously been found in some of the tested regions during goal-directed behaviour (Gruber et al., 2009). Our analysis highlights the encoding of reward information above 40Hz that may have previously been unexplored. This predictive disparity suggests frequencies below 40 Hz, while containing relatively high-powered reinforcement outcome information, encode redundant reinforcement data that may be ascertained from the higher frequencies, whereas the reinforcement information of the higher frequencies seems to be exclusive. Alternatively, both of these ranges may contain mutually exclusive reinforcement information such that information encoded above 40Hz is more predictive of the reinforcement outcome. We speculate the former is more likely because past evidence has shown alteration of gamma activity following reward in the VST (Berke, 2009; van der Meer, 2009). Specifically, in regards to reinforcement outcome, Berke (2009) showed a significant power increase in the 70-90 Hz range in unrewarded trials in contrast to rewarded trials.

The results of our research introduce more interesting questions: which sensory systems are responding to reinforcement? Is the animal conditioned to the click of the solenoid by which the reward is delivered? Is the animal conditioned to the change in lighting that occurs if the incorrect well is chosen? Is an olfactory/gustatory response the first to trigger the widespread reinforcement signal? We suspect that the animal uses multiple sensory modalities to recognize the presence, or absence, of reward; this could contribute to the ubiquitous nature of the reinforcement encoding. Sensory processing seems to be relatively widespread throughout the regions from which we recorded. Olfactory information is projected, either directly or indirectly, to the amygdala, hippocampus, OFC and VST (Paxinos, 2004). Other sensory information is likely

distributed throughout these regions, via thalamic projections to the striatum, cortex, BLA and hippocampus (Carleton, Accolla, & Simon, 2010; Su & Bentivoglio, 1990; Voorn et al., 2004). However, although it may be reasonable to suggest pervasive parallel processing of the sensory information received at the time of reinforcement, our data suggest differential encoding of the *valence* of the reinforcement, which may stem from differential encoding of sensory inputs in the recorded regions.

Machine learning has recently emerged as a promising tool for the analysis of neuroscience data, specifically due to advancements in explainable artificial intelligence (Buzsáki et al., 2018). We have shown application of these techniques to field potentials from rodents can elucidate information encoding. Our results suggest degeneracy within the decision-making circuit: these structurally distinct regions all encode reinforcement (Edelman & Gally, 2001). This degeneracy is likely evolutionarily advantageous, enabling the animal to maintain this fundamental system despite disease or damage (Noppeney, Friston, & Price, 2004).

5. General Discussion

This thesis has presented two different experiments that each utilized the same choice task. The first, examining sex differences, highlighted differential behaviour of female rats outside of the task context. I presented evidence to suggest this disparity was due to differences in exploratory drive. However, within the task context, there was no significant sex difference. The latter study analyzed field potentials in multiple structures in the cortico-striatal-limbic circuit as male rats completed the choice task. From the results of this study, I have suggested the outcome aspect of reinforcement processing is distributed throughout all of the recorded brain regions. The second experiment helps to inform the negative finding of the first. If the reward signal is distributed throughout the decision-making system, it is not surprising that neurobiological sexual dimorphisms within one structure, such as the dopaminergic differences in the striatum (Becker et al., 2012; Mozley et al., 2001; Walker et al., 2006, 2000), have little effect on the performance and behaviour of the animals within the context of the task.

Throughout this thesis I have speculated on the influence of dopamine (DA) on the obtained results, although it was never directly measured. The role of DA I described is that which was proposed by Schultz of prediction error: DA neurons exhibit a phasic response following an *unexpected* reward, and their activity is depressed following an *unexpected* omission of reward (Schultz et al., 1997). As previously noted, the reward in this task is primarily stochastic, and the rat should have no expectancy of reward in each trial. The rats lick at a similar rate prior to reward delivery or omission, suggesting there is some inherent expectancy. However, we have done no direct experimentation or analysis to investigate this expectancy. Thus, more research is required to investigate the

mechanism driving the observed ubiquity of reinforcement processing. This finding may be researched further via recording of dopaminergic midbrain neurons simultaneously with recording of the structures investigated here during the choice task. The ubiquity may also be researched via direct activation of these DA neurons and recording of the downstream effects. The latter, via optogenetic stimulation, has recently garnered profound results. Activation of midbrain DA neurons has been shown to initiate locomotion and drive operant conditioning (Howe & Dombeck, 2016; Kim et al., 2012; Schultz, Stauffer, & Lak, 2017).

In many cases, human research investigating the neurobiological correlates and treatments for clinical decision-making disorders has suggested distributed dysfunction. For example, major depressive disorder is characterized by an impairment in the typical response to reward, such that a person may not experience pleasure from activities they previously enjoyed (Whitton, Treadway, & Pizzagalli, 2015). Among others, dysfunction of the OFC, mPFC, BLA, HPC, and the striatum have each been linked to depression (Pandya, Altinay, Malone, Anand, & Anand, 2012; Rigucci, Serafini, Pompili, Kotzalidis, & Tatarelli, 2010; Schlaepfer et al., 2008). The widespread dysfunction associated with this condition is consistent with the results obtained from rats in this thesis. Further research investigating the mechanism behind this ubiquity in reward processing may help inform future study of disordered decision-making.

6. References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., … Brain, G. (2016). TensorFlow: A system for large-scale machine learning. *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*, 265–283. Retrieved from https://tensorflow.org.

Alstott, J., & Timberlake, W. (2009). Effects of rat sex differences and lighting on locomotor exploration of a circular open field with free-standing central corners and without peripheral walls. *Behavioural Brain Research*, *196*(2), 214–219. https://doi.org/10.1016/j.bbr.2008.09.001

Ambroggi, F., Ishikawa, A., Fields, H. L., & Nicola, S. M. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron*, *59*(4), 648–661. https://doi.org/10.1016/j.neuron.2008.07.004

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*(1), 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709

Atallah, H. E., McCool, A. D., Howe, M. W., & Graybiel, A. M. (2014). Neurons in the ventral striatum exhibit cell-type-specific representations of outcome during learning. *Neuron*, *82*(5), 1145–1156. https://doi.org/10.1016/j.neuron.2014.04.021

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, *35*(1), 48–69. https://doi.org/10.1038/npp.2009.131

Barker, G. R. I., Bird, F., Alexander, V., & Warburton, E. C. (2007). Recognition memory for objects, place, and temporal order: A disconnection analysis of the role of the medial prefrontal cortex and perirhinal cortex. *Journal of Neuroscience*, *27*(11), 2948–2957. https://doi.org/10.1523/JNEUROSCI.5289-06.2007

Barnett, L., & Seth, A. K. (2014). The mvgc multivariate granger causality toolbox: A new approach to granger-causal inference. *Journal of Neuroscience Methods*, *223*, 50–68. https://doi.org/10.1016/j.jneumeth.2013.10.018

Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, *47*(1), 129–141. https://doi.org/10.1016/j.neuron.2005.05.020

Becker, J. B., Perry, A. N., & Westenbroek, C. (2012). Sex differences in the neural mechanisms mediating addiction: A new synthesis and hypothesis. *Biology of Sex Differences*, *3*(1), 14. https://doi.org/10.1186/2042-6410-3-14

Berke, J. D. (2009). Fast oscillations in cortical-striatal networks switch frequency following rewarding events and stimulant drugs. *European Journal of Neuroscience*,

30(5), 848–859. https://doi.org/10.1111/j.1460-9568.2009.06843.x

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28*(3), 309–369. https://doi.org/10.1016/S0165-0173(98)00019-8

Berridge, K. C., & Whishaw, I. Q. (1992). Cortex, striatum and cerebellum: Control of serial order in a grooming sequence. *Exp Brain Res*, *90*, 275–290. Retrieved from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/46562/221_2004_Article_BF00227239.pdf?sequence=1&isAllowed=y

Björklund, A., & Dunnett, S. B. (2007). Dopamine neuron systems in the brain: An update. *Trends in Neurosciences*, *30*(5), 194–202. https://doi.org/10.1016/j.tins.2007.03.006

Bressler, S. L., & Seth, A. K. (2011). Wiener–Granger causality: A well established methodology. *NeuroImage*, *58*(2), 323–329. https://doi.org/10.1016/J.NEUROIMAGE.2010.02.059

Bromberg-Martin, E. S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: Rewarding, aversive, and alerting. *Neuron*, *68*(5), 815–834. https://doi.org/10.1016/J.NEURON.2010.11.022

Brunello, N., Masotto, C., Steardo, L., Markstein, R., & Racagni, G. (1995). New insights into the biology of schizophrenia through the mechanism of action of clozapine. *Neuropsychopharmacology*, *13*(3), 177–213. https://doi.org/10.1016/0893-133X(95)00068-O

Buerger, A., Rozhitskaya, O., Sherwood, M. C., Dorfman, A. L., Bisping, E., Abel, E. D., … Jay, P. Y. (2006). Dilated cardiomyopathy resulting from high-level myocardial expression of cre-recombinase. *Journal of Cardiac Failure*, *12*(5), 392–398. https://doi.org/10.1016/j.cardfail.2006.03.002

Burton, A. C., Nakamura, K., & Roesch, M. R. (2015). From ventral-medial to dorsal-lateral striatum: Neural correlates of reward-guided decision-making. *Neurobiology of Learning and Memory*, *117*, 51–59. https://doi.org/10.1016/j.nlm.2014.05.003

Burton, C. L., & Fletcher, P. J. (2012). Age and sex differences in impulsive action in rats: The role of dopamine and glutamate. *Behavioural Brain Research*, *230*(1), 21–33. https://doi.org/10.1016/j.bbr.2012.01.046

Buzsáki, G., Dzirasa, K., Sapiro, G., Rudin, C., Ba, D., Widge, A. S., … Heller, K. (2018). A shared vision for machine learning in neuroscience. *The Journal of Neuroscience*, *38*(7), 1601–1607. https://doi.org/10.1523/jneurosci.0508-17.2018

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367–383. https://doi.org/10.1037/0033-2909.125.3.367

Carleton, A., Accolla, R., & Simon, S. A. (2010). Coding in the mammalian gustatory system. *Trends in Neurosciences*, *33*(7), 326–334. https://doi.org/10.1016/j.tins.2010.04.002

Caswell, T. A., Droettboom, M., Hunter, J., Firing, E., Lee, A., Stansby, D., … Würtz, P. (2018). *Matplotlib*. https://doi.org/10.5281/ZENODO.1482099

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, *22*(11), 2575–2586. https://doi.org/10.1093/cercor/bhr332

Chollet, F. (2015). *Keras*. Retrieved from https://keras.io/

Coutureau, E., Esclassan, F., Di Scala, G., & Marchand, A. R. (2012). The role of the rat medial prefrontal cortex in adapting to changes in instrumental contingency. *PLoS ONE*, *7*(4), e33302. https://doi.org/10.1371/journal.pone.0033302

Cross, C. P., Copping, L. T., & Campbell, A. (2011). Sex differences in impulsivity: A meta-analysis. *Psychological Bulletin*, *137*(1), 97–130. https://doi.org/10.1037/a0021591

D'Errico, J. (2006). *Inpaint_nans*. Retrieved from https://www.mathworks.com/matlabcentral/fileexchange/4551-inpaint_nans

Datla, K. P., Murray, H. E., Pillai, A. V, Gillies, G. E., & Dexter, D. T. (2003). Differences in dopaminergic neuroprotective effects of estrogen during estrous cycle. *Neuroreport*, *14*(1), 47–50. https://doi.org/10.1097/01.wnr.0000050300.92401.45

Dauer, W., & Przedborski, S. (2003). Parkinson's disease: Mechanisms and models. *Neuron*, *39*(6), 889–909. https://doi.org/10.1016/S0896-6273(03)00568-3

Devan, B. D., Hong, N. S., & McDonald, R. J. (2011, September). Parallel associative processing in the dorsal striatum: Segregation of stimulus-response and cognitive control subregions. *Neurobiology of Learning and Memory*, Vol. 96, pp. 95–120. https://doi.org/10.1016/j.nlm.2011.06.002

Donovan, C. H., Wong, S. A., Randolph, S. H., Stark, R. A., Gibb, R. L., & Gruber, A. J. (2018). Sex differences in rat decision-making: The confounding role of extraneous feeder sampling between trials. *Behavioural Brain Research*, *342*, 62–69. https://doi.org/10.1016/j.bbr.2018.01.018

Doya, K. (2008, April 1). Modulators of decision making. *Nature Neuroscience*, Vol. 11, pp. 410–416. https://doi.org/10.1038/nn2077

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(24), 13763–13768. https://doi.org/10.1073/pnas.231499798

Eubig, P. A., Noe, T. E., Floresco, S. B., Sable, J. J., & Schantz, S. L. (2014). Sex differences in response to amphetamine in adult long-evans rats performing a delay-discounting task. *Pharmacology Biochemistry and Behavior*, *118*, 1–9. https://doi.org/10.1016/j.pbb.2013.12.021

Euston, D. R., Gruber, A. J., & McNaughton, B. L. (2012, December 20). The role of medial prefrontal cortex in memory and decision making. *Neuron*, Vol. 76, pp. 1057–1070. https://doi.org/10.1016/j.neuron.2012.12.002

Everitt, B. J., & Robbins, T. W. (2005, November 1). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*, Vol. 8, pp. 1481–1489. https://doi.org/10.1038/nn1579

Flagel, S. B., Watson, S. J., Robinson, T. E., & Akil, H. (2007). Individual differences in the propensity to approach signals vs goals promote different adaptations in the dopamine system of rats. *Psychopharmacology*, *191*(3), 599–607. https://doi.org/10.1007/s00213-006-0535-8

Floresco, S. B., St. Onge, J. R., Ghods-Sharifi, S., & Winstanley, C. A. (2008). Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cognitive, Affective and Behavioral Neuroscience*, *8*(4), 375–389. https://doi.org/10.3758/CABN.8.4.375

Forni, P. E., Scuoppo, C., Imayoshi, I., Taulli, R., Dastrù, W., Sala, V., … Ponzetto, C. (2006). High levels of cre expression in neuronal progenitors cause defects in brain development leading to microencephaly and hydrocephaly. *Journal of Neuroscience*, *26*(37), 9593–9602. https://doi.org/10.1523/JNEUROSCI.2815-06.2006

Gayle, D. A., Desai, M., Casillas, E., Beloosesky, R., & Ross, M. G. (2006). Gender-specific orexigenic and anorexigenic mechanisms in rats. *Life Sciences*, *79*(16), 1531–1536. https://doi.org/10.1016/j.lfs.2006.04.015

Gruber, A. J., Hussain, R. J., & O'Donnell, P. (2009). The nucleus accumbens: A switchboard for goal-directed behaviors. *PLoS ONE*, *4*(4), e5062. https://doi.org/10.1371/journal.pone.0005062

Gruber, A. J., & McDonald, R. J. (2012). Context, emotion, and the strategic pursuit of goals: Interactions among multiple brain systems controlling motivated behavior. *Frontiers in Behavioral Neuroscience*, *6*, 50. https://doi.org/10.3389/fnbeh.2012.00050

Gruber, A. J., & Thapa, R. (2016). The memory trace supporting lose-shift responding decays rapidly after reward omission and is distinct from other learning mechanisms in rats. *ENeuro*, *3*(6). https://doi.org/10.1523/ENEURO.0167-16.2016

Gruber, A. J., Thapa, R., & Randolph, S. H. (2017). Feeder approach between trials is increased by uncertainty and affects subsequent choices. *Eneuro*, *4*(6), ENEURO.0437-17.2017. https://doi.org/10.1523/ENEURO.0437-17.2017

Haber, S. N. (2003). The primate basal ganglia: Parallel and integrative networks. *Journal of Chemical Neuroanatomy*, *26*(4), 317–330. https://doi.org/10.1016/j.jchemneu.2003.10.003

HajiHosseini, A., & Holroyd, C. B. (2015). Sensitivity of frontal beta oscillations to reward valence but not probability. *Neuroscience Letters*, *602*, 99–103. https://doi.org/10.1016/J.NEULET.2015.06.054

HajiHosseini, A., Rodríguez-Fornells, A., & Marco-Pallarés, J. (2012). The role of beta-gamma oscillations in unexpected rewards processing. *NeuroImage*, *60*(3), 1678–1685. https://doi.org/10.1016/J.NEUROIMAGE.2012.01.125

Hall, B. D., & Willink, R. (2001). Does "welch-satterthwaite" make a good uncertainty estimate? *Metrologia*, *38*(1), 9–15. https://doi.org/10.1088/0026-1394/38/1/2

Hammerslag, L. R., & Gulley, J. M. (2014). Age and sex differences in reward behavior in adolescent and adult rats. *Developmental Psychobiology*, *56*(4), 611–621. https://doi.org/10.1002/dev.21127

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Herreras, O. (2016). Local field potentials: Myths and misunderstandings. *Frontiers in Neural Circuits*, *10*, 101. https://doi.org/10.3389/fncir.2016.00101

Hitchcott, P. K., Quinn, J. J., & Taylor, J. R. (2007). Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cerebral Cortex*, *17*(12), 2820–2827. https://doi.org/10.1093/cercor/bhm010

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, *1*(4), 304–309. https://doi.org/10.1038/1124

Horvitz, J. C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*, *96*(4), 651–656. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10727783

Howe, M. W., & Dombeck, D. A. (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature*, *535*(7613), 505–510. https://doi.org/10.1038/nature18942

Hussain, R., Tassabehji, M., Ashton, N., & Glazier, J. (2017). A rat model of maternal intermittent fasting during pregnancy impairs fetal growth but does not alter maternofetal folate transport. *Placenta*, *57*, 286. https://doi.org/10.1016/j.placenta.2017.07.201

Ito, M., & Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *Journal of Neuroscience*, *29*(31), 9861–

9874. https://doi.org/10.1523/JNEUROSCI.6157-08.2009

Jentsch, J. D., & Taylor, J. R. (2003). Sex-related differences in spatial divided attention and motor impulsivity in rats. *Behavioral Neuroscience*, *117*(1), 76–83. https://doi.org/10.1037/0735-7044.117.1.76

Johnston, A. L., & File, S. E. (1991). Sex differences in animal tests of anxiety. *Physiology and Behavior*, *49*(2), 245–250. https://doi.org/10.1016/0031-9384(91)90039-Q

Kearns, D. N., Gomez-Serrano, M. A., Weiss, S. J., & Riley, A. L. (2006). A comparison of lewis and fischer rat strains on autoshaping (sign-tracking), discrimination reversal learning and negative automaintenance. *Behavioural Brain Research*, *169*(2), 193–200. https://doi.org/10.1016/j.bbr.2006.01.005

Kim, H., Lee, D., & Jung, M. W. (2013). Signals for previous goal choice persist in the dorsomedial, but not dorsolateral striatum of rats. *Journal of Neuroscience*, *33*(1), 52–63. https://doi.org/10.1523/JNEUROSCI.2422-12.2013

Kim, K. M., Baratta, M. V., Yang, A., Lee, D., Boyden, E. S., & Fiorillo, C. D. (2012). Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. *PLoS ONE*, *7*(4), e33612. https://doi.org/10.1371/journal.pone.0033612

Kobayashi, S., Nomoto, K., Watanabe, M., Hikosaka, O., Schultz, W., & Sakagami, M. (2006). Influences of rewarding and aversive outcomes on activity in macaque lateral prefrontal cortex. *Neuron*, *51*(6), 861–870. https://doi.org/10.1016/J.NEURON.2006.08.031

Kolb, B., & Cioe, J. (1996). Sex-related differences in cortical function after medial frontal lesions in rats. *Behavioral Neuroscience*, *110*(6), 1271–1281. https://doi.org/10.1037/0735-7044.110.6.1271

Kolb, B., & Stewart, J. (1991). Sex-Related differences in dendritic branching of cells in the prefrontal cortex of rats. *Journal of Neuroendocrinology*, *3*(1), 95–99. https://doi.org/10.1111/j.1365-2826.1991.tb00245.x

Korol, D. L., Malin, E. L., Borden, K. A., Busby, R. A., & Couper-Leo, J. (2004). Shifts in preferred learning strategy across the estrous cycle in female rats. *Hormones and Behavior*, *45*(5), 330–338. https://doi.org/10.1016/j.yhbeh.2004.01.005

Kotikalapudi, R. (2017). *Keras-vis*. Retrieved from https://github.com/raghakot/keras-vis

Lammers, C. H., D'Souza, U., Qin, Z. H., Lee, S. H., Yajima, S., & Mouradian, M. M. (1999). Regulation of striatal dopamine receptors by estrogen. *Synapse*, *34*(3), 222–227. https://doi.org/10.1002/(SICI)1098-2396(19991201)34:3<222::AID-SYN6>3.0.CO;2-J

Lau, B., & Glimcher, P. W. (2007). Action and outcome encoding in the primate caudate

nucleus. *Journal of Neuroscience*, *27*(52), 14502–14514. https://doi.org/10.1523/jneurosci.3060-07.2007

Lee, D., Conroy, M. L., McGreevy, B. P., & Barraclough, D. J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cognitive Brain Research*, *22*(1), 45–58. https://doi.org/10.1016/j.cogbrainres.2004.07.007

Loonstra, A., Vooijs, M., Beverloo, H. B., Allak, B. A., van Drunen, E., Kanaar, R., … Jonkers, J. (2001). Growth inhibition and dna damage induced by cre recombinase in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(16), 9209–9214. https://doi.org/10.1073/pnas.161269798

Lukkes, J. L., Thompson, B. S., Freund, N., & Andersen, S. L. (2016). The developmental inter-relationships between activity, novelty preferences, and delay discounting in male and female rats. *Developmental Psychobiology*, *58*(2), 231–242. https://doi.org/10.1002/dev.21368

Lynn, D. A., & Brown, G. R. (2009). The ontogeny of exploratory behavior in male and female adolescent rats (rattus norvegicus). *Developmental Psychobiology*, *51*(6), 513–520. https://doi.org/10.1002/dev.20386

Marco-Pallares, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T. F., & Rodríguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia*, *46*(1), 241–248. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2007.07.016

Marco-Pallarés, J., Münte, T. F., & Rodríguez-Fornells, A. (2015). The role of high-frequency oscillatory activity in reward processing and learning. *Neuroscience & Biobehavioral Reviews*, *49*, 1–7. https://doi.org/10.1016/J.NEUBIOREV.2014.11.014

Markham, J. A., & Juraska, J. M. (2002). Aging and sex influence the anatomy of the rat anterior cingulate cortex. *Neurobiology of Aging*, *23*(4), 579–588. https://doi.org/10.1016/S0197-4580(02)00004-0

Mas-Herrero, E., & Marco-Pallarés, J. (2014). Frontal theta oscillatory activity is a common mechanism for the computation of unexpected outcomes and learning rate. *Journal of Cognitive Neuroscience*, *26*(3), 447–458. https://doi.org/10.1162/jocn_a_00516

Mashhoori, A., Hashemnia, S., McNaughton, B. L., Euston, D. R., & Gruber, A. J. (2018). Rat anterior cingulate cortex recalls features of remote reward locations after disfavoured reinforcements. *ELife*, *7*. https://doi.org/10.7554/eLife.29793

McCracken, C. B., & Grace, A. A. (2007). High-frequency deep brain stimulation of the nucleus accumbens region suppresses neuronal activity and selectively modulates afferent drive in rat orbitofrontal cortex in vivo. *Journal of Neuroscience*, *27*(46), 12601–12610. https://doi.org/10.1523/JNEUROSCI.3750-07.2007

Mendez, M. F. (2017). A neurology of the conservative-liberal dimension of political ideology. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *29*(2), 86–94. https://doi.org/10.1176/appi.neuropsych.16030051

Meyer, P. J., Lovic, V., Saunders, B. T., Yager, L. M., Flagel, S. B., Morrow, J. D., & Robinson, T. E. (2012). Quantifying individual variation in the propensity to attribute incentive salience to reward cues. *PLoS ONE*, *7*(6), e38987. https://doi.org/10.1371/journal.pone.0038987

Mitra, P. (2007). *Observed brain dynamics*. Retrieved from http://chronux.org/

Morrison, S. E., & Nicola, S. M. (2014). Neurons in the nucleus accumbens promote selection bias for nearer objects. *Journal of Neuroscience*, *34*(42), 14147–14162. https://doi.org/10.1523/JNEUROSCI.2197-14.2014

Mozley, L. H., Gur, R. C., Mozley, P. D., & Gur, R. E. (2001). Striatal dopamine transporters and cognitive functioning in healthy men and women. *American Journal of Psychiatry*, *158*(9), 1492–1499. https://doi.org/10.1176/appi.ajp.158.9.1492

Murphy, M. J. M., & Deutch, A. Y. (2018). Organization of afferents to the orbitofrontal cortex in the rat. *Journal of Comparative Neurology*, *526*(9), 1498–1526. https://doi.org/10.1002/cne.24424

Murray, E. A., O'Doherty, J. P., & Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *Journal of Neuroscience*, *27*(31), 8166–8169. https://doi.org/10.1523/JNEUROSCI.1556-07.2007

Nasello, A. G., MacHado, C., Bastos, J. F., & Felicio, L. F. (1998). Sudden darkness induces a high activity-low anxiety state in male and female rats. *Physiology and Behavior*, *63*(3), 451–454. https://doi.org/10.1016/S0031-9384(97)00462-9

Noppeney, U., Friston, K. J., & Price, C. J. (2004). Degenerate neuronal systems sustaining cognitive functions. *Journal of Anatomy*, *205*(6), 433–442. https://doi.org/10.1111/j.0021-8782.2004.00343.x

Orsini, C. A., & Setlow, B. (2017). Sex differences in animal models of decision making. *Journal of Neuroscience Research*, *95*(1–2), 260–269. https://doi.org/10.1002/jnr.23810

Orsini, C. A., Willis, M. L., Gilbert, R. J., Bizon, J. L., & Setlow, B. (2016). Sex differences in a rat model of risky decision making. *Behavioral Neuroscience*, *130*(1), 50–61. https://doi.org/10.1037/bne0000111

Packard, M. G., & White, N. M. (1991). Dissociation of hippocampus and caudate nucleus memory systems by posttraining intracerebral injection of dopamine agonists. *Behavioral Neuroscience*, *105*(2), 295–306. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1675062

Pandya, M., Altinay, M., Malone, D. A., Anand, A., & Anand, A. (2012). Where in the brain is depression? *Current Psychiatry Reports*, *14*(6), 634–642. https://doi.org/10.1007/s11920-012-0322-7

Pasqualini, C., Olivier, V., Guibert, B., Frain, O., & Leviel, V. (2002). Acute stimulatory effect of estradiol on striatal dopamine synthesis. *Journal of Neurochemistry*, *65*(4), 1651–1657. https://doi.org/10.1046/j.1471-4159.1995.65041651.x

Paxinos, G. (2004). The rat nervous system. In G. Paxinos (Ed.), *The Rat Nervous System* (3rd ed.). https://doi.org/10.1016/B978-012547638-6/50027-4

Paxinos, G., & Watson, C. (2005). *The rat brain in stereotaxic coordinates* (5th ed.). San Diego: Elsevier Academic Press.

Peak, J. N., Turner, K. M., & Burne, T. H. J. (2015). The effect of developmental vitamin d deficiency in male and female sprague-dawley rats on decision-making using a rodent gambling task. *Physiology and Behavior*, *138*, 319–324. https://doi.org/10.1016/j.physbeh.2014.09.007

Pennartz, C. M. A., Ito, R., Verschure, P. F. M. J., Battaglia, F. P., & Robbins, T. W. (2011, October). The hippocampal-striatal axis in learning, prediction and goal-directed behavior. *Trends in Neurosciences*, Vol. 34, pp. 548–559. https://doi.org/10.1016/j.tins.2011.08.001

Phillips, G. D., Salussolia, E., & Hitchcott, P. K. (2010). Role of the mesoamygdaloid dopamine projection in emotional learning. *Psychopharmacology*, *210*(3), 303–316. https://doi.org/10.1007/s00213-010-1813-z

Pitchers, K. K., Flagel, S. B., O'Donnell, E. G., Solberg Woods, L. C., Sarter, M., & Robinson, T. E. (2015). Individual variation in the propensity to attribute incentive salience to a food cue: Influence of sex. *Behavioural Brain Research*, *278*, 462–469. https://doi.org/10.1016/j.bbr.2014.10.036

Pratt, W. E., & Mizumori, S. J. . (2001). Neurons in rat medial prefrontal cortex show anticipatory rate changes to predictable differential rewards in a spatial memory task. *Behavioural Brain Research*, *123*(2), 165–183. https://doi.org/10.1016/S0166-4328(01)00204-2

Quilodran, R., Rothé, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*(2), 314–325. https://doi.org/10.1016/j.neuron.2007.11.031

Ray, J., & Hansen, S. (2004). Temperament in the rat: Sex differences and hormonal influences on harm avoidance and novelty seeking. *Behavioral Neuroscience*, *118*(3), 488–497. https://doi.org/10.1037/0735-7044.118.3.488

Reig, R., & Silberberg, G. (2014). Multisensory integration in the mouse striatum. *Neuron*, *83*(5), 1200–1212. https://doi.org/10.1016/j.neuron.2014.07.033

Rigucci, S., Serafini, G., Pompili, M., Kotzalidis, G. D., & Tatarelli, R. (2010). Anatomical and functional correlates in major depressive disorder: The contribution of neuroimaging studies. *The World Journal of Biological Psychiatry*, *11*(2–2), 165–180. https://doi.org/10.3109/15622970903131571

Robinson, T. E., Yager, L. M., Cogan, E. S., & Saunders, B. T. (2014, January 1). On the motivational properties of reward cues: Individual differences. *Neuropharmacology*, Vol. 76, pp. 450–459. https://doi.org/10.1016/j.neuropharm.2013.05.040

Roesch, M. R., Calu, D. J., Esber, G. R., & Schoenbaum, G. (2010). Neural correlates of variations in event processing during learning in basolateral amygdala. *Journal of Neuroscience*, *30*(7), 2464–2471. https://doi.org/10.1523/JNEUROSCI.5781-09.2010

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., & Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *Journal of Neuroscience*, *29*(42), 13365–13376. https://doi.org/10.1523/JNEUROSCI.2572-09.2009

Roesch, M. R., Taylor, A. R., & Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron*, *51*(4), 509–520. https://doi.org/10.1016/j.neuron.2006.06.027

Rorden, C., & Karnath, H.-O. (2004). Using human brain lesions to infer function: A relic from a past era in the fmri age? *Nature Reviews Neuroscience*, *5*(10), 812–819. https://doi.org/10.1038/nrn1521

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562. https://doi.org/10.1038/nrn3776

Schilman, E. A., Uylings, H. B. M., Graaf, Y. G., Joel, D., & Groenewegen, H. J. (2008). The orbital cortex in rats topographically projects to central parts of the caudate–putamen complex. *Neuroscience Letters*, *432*(1), 40–45. https://doi.org/10.1016/J.NEULET.2007.12.024

Schlaepfer, T. E., Cohen, M. X., Frick, C., Kosel, M., Brodesser, D., Axmacher, N., … Sturm, V. (2008). Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. *Neuropsychopharmacology*, *33*, 368–377. https://doi.org/10.1038/sj.npp.1301408

Schmidt, E. E., Taylor, D. S., Prigge, J. R., Barnett, S., & Capecchi, M. R. (2000). Illegitimate cre-dependent chromosome rearrangements in transgenic mouse spermatids. *Proceedings of the National Academy of Sciences*, *97*(25), 13702–13707. https://doi.org/10.1073/pnas.240471297

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nature*

Reviews Neuroscience, *10*(12), 885–892. https://doi.org/10.1038/nrn2753

Schoenbaum, G., Setlow, B., Saddoris, M. P., & Gallagher, M. (2003). Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron*, *39*(5), 855–867. https://doi.org/10.1016/S0896-6273(03)00474-4

Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience*, *17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306). Retrieved from http://science.sciencemag.org/content/275/5306/1593.full

Schultz, W., Stauffer, W. R., & Lak, A. (2017). The phasic dopamine signal maturing: From reward via behavioural activation to formal economic utility. *Current Opinion in Neurobiology*, *43*, 139–148. https://doi.org/10.1016/J.CONB.2017.03.013

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision*, *2017–Octob*, 618–626. https://doi.org/10.1109/ICCV.2017.74

Shenhav, A., Cohen, J. D., & Botvinick, M. M. (2016). Dorsal anterior cingulate cortex and the value of control. *Nature Neuroscience*, *19*(10), 1286–1291. https://doi.org/10.1038/nn.4384

Shin, J. H., Kim, D., & Jung, M. W. (2018). Differential coding of reward and movement information in the dorsomedial striatal direct and indirect pathways. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02817-1

Simmons, J. M., Ravel, S., Shidara, M., & Richmond, B. J. (2007). A comparison of reward-contingent neuronal activity in monkey orbitofrontal cortex and ventral striatum: Guiding actions toward rewards. *Annals of the New York Academy of Sciences*, *1121*(1), 376–394. https://doi.org/10.1196/annals.1401.028

Skelin, I., Hakstol, R., VanOyen, J., Mudiayi, D., Molina, L. A., Holec, V., … Gruber, A. J. (2014). Lesions of dorsal striatum eliminate lose-switch responding but not mixed-response strategies in rats. *European Journal of Neuroscience*, *39*(10), 1655–1663. https://doi.org/10.1111/ejn.12518

Skelin, I., Needham, M. A., Molina, L. M., Metz, G. A. S., & Gruber, A. J. (2015). Multigenerational prenatal stress increases the coherence of brain signaling among cortico–striatal–limbic circuits in adult rats. *Neuroscience*, *289*, 270–278. https://doi.org/10.1016/J.NEUROSCIENCE.2015.01.009

Smethells, J. R., Swalve, N. L., Eberly, L. E., & Carroll, M. E. (2016). Sex differences in the reduction of impulsive choice (delay discounting) for cocaine in rats with

atomoxetine and progesterone. *Psychopharmacology*, *233*(15–16), 2999–3008. https://doi.org/10.1007/s00213-016-4345-3

Solanto, M. V. (1998). Neuropsychopharmacological mechanisms of stimulant drug action in attention-deficit hyperactivity disorder: A review and integration. *Behavioural Brain Research*, *94*(1), 127–152. https://doi.org/10.1016/S0166-4328(97)00175-7

Stalnaker, T. A., Calhoon, G. G., Ogawa, M., Roesch, M. R., & Schoenbaum, G. (2012). Reward prediction error signaling in posterior dorsomedial striatum is action specific. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(30), 10296–10305. https://doi.org/10.1523/JNEUROSCI.0832-12.2012

Su, H.-S., & Bentivoglio, M. (1990). Thalamic midline cell populations projecting to the nucleus accumbens, amygdala, and hippocampus in the rat. *The Journal of Comparative Neurology*, *297*(4), 582–593. https://doi.org/10.1002/cne.902970410

Sugase-Miyamoto, Y., & Richmond, B. (2005). Neuronal signals in the monkey basolateral amygdala during reward schedules. *Journal of Neuroscience*, *25*(48), 11071–11083. https://doi.org/10.1523/JNEUROSCI.1796-05.2005

Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron*, *66*(3), 449–460. https://doi.org/10.1016/j.neuron.2010.03.033

Sutcliffe, J. S., Marshall, K. M., & Neill, J. C. (2007). Influence of gender on working and spatial memory in the novel object recognition task in the rat. *Behavioural Brain Research*, *177*(1), 117–125. https://doi.org/10.1016/j.bbr.2006.10.029

The MathWorks Inc. (2018). *MATLAB* (9.5.0). 9.5.0. Natick, Massachusetts: The MathWorks Inc.,.

Torres-Hernández, A. R., & González-Vegas, J. A. (2005). Effects of 17β-estradiol on the spontaneous activity of substantia nigra neurons: Evidence for a non-genomic mechanism. *Brain Research*, *1049*(1), 1–7. https://doi.org/10.1016/j.brainres.2005.04.085

Tropp, J., & Markus, E. J. (2001). Sex differences in the dynamics of cue utilization and exploratory behavior. *Behavioural Brain Research*, *119*(2), 143–154. https://doi.org/10.1016/S0166-4328(00)00345-4

Van den Bos, R., Homberg, J., & de Visser, L. (2013). A critical review of sex differences in decision-making tasks: Focus on the iowa gambling task. *Behavioural Brain Research*, Vol. 238, pp. 95–108. https://doi.org/10.1016/j.bbr.2012.10.002

Van den Bos, R., Jolles, J., Van der Knaap, L., Baars, A., & De Visser, L. (2012). Male and female wistar rats differ in decision-making performance in a rodent version of the iowa gambling task. *Behavioural Brain Research*, *234*(2), 375–379.

https://doi.org/10.1016/j.bbr.2012.07.015

Van den Bos, R., Lasthuis, W., den Heijer, E., Van der Harst, J., & Spruijt, B. (2006). Toward a rodent model of the iowa gambling task. *Behavior Research Methods*, *38*(3), 470–478. https://doi.org/10.3758/BF03192801

van der Meer, M. A. A. (2009). Low and high gamma oscillations in rat ventral striatum have distinct relationships to behavior, reward, and spiking activity on a learned spatial decision task. *Frontiers in Integrative Neuroscience*, *3*, 9. https://doi.org/10.3389/neuro.07.009.2009

van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C., & Redish, A. D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron*, *67*(1), 25–32. https://doi.org/10.1016/j.neuron.2010.06.023

van Hasselt, F. N., de Visser, L., Tieskens, J. M., Cornelisse, S., Baars, A. M., Lavrijsen, M., … Joëls, M. (2012). Individual variations in maternal care early in life correlate with later life decision-making and c-fos expression in prefrontal subregions of rats. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0037820

Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron*, *72*(1), 166–177. https://doi.org/10.1016/j.neuron.2011.08.011

Voorn, P., Vanderschuren, L. J. M. ., Groenewegen, H. J., Robbins, T. W., & Pennartz, C. M. . (2004). Putting a spin on the dorsal–ventral divide of the striatum. *Trends in Neurosciences*, *27*(8), 468–474. https://doi.org/10.1016/J.TINS.2004.06.006

Wade, G. N. (1972, March). Gonadal hormones and behavioral regulation of body weight. *Physiology and Behavior*, Vol. 8, pp. 523–534. https://doi.org/10.1016/0031-9384(72)90340-X

Wade, G. N., Schneider, J. E., & Li, H.-Y. (1996). Control of fertility by metabolic cues. *The American Journal of Physiology*, *270*(1), E1–E19. https://doi.org/10.1152/ajpendo.1996.270.1.E1

Walker, Q. D., Ray, R., & Kuhn, C. M. (2006). Sex differences in neurochemical effects of dopaminergic drugs in rat striatum. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *31*(6), 1193–1202. https://doi.org/10.1038/sj.npp.1300915

Walker, Q. D., Rooney, M. B., Wightman, R. M., & Kuhn, C. M. (2000). Dopamine release and uptake are greater in female than male rat striatum as measured by fast cyclic voltammetry. *Neuroscience*, *95*(4), 1061–1070. https://doi.org/10.1016/S0306-4522(99)00500-X

Wang, A. Y., Miura, K., & Uchida, N. (2013). The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nature Neuroscience*,

16(5), 639–647. https://doi.org/10.1038/nn.3377

Weible, A. P., Rowland, D. C., Pang, R., & Kentros, C. (2009). Neural correlates of novel object and novel location recognition behavior in the mouse anterior cingulate cortex. *Journal of Neurophysiology*, *102*(4), 2055–2068. https://doi.org/10.1152/jn.00214.2009

Whitton, A. E., Treadway, M. T., & Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Current Opinion in Psychiatry*, *28*(1), 7–12. https://doi.org/10.1097/YCO.0000000000000122

Wingerden, M. van, Vinck, M., Lankelma, J. V., & Pennartz, C. M. A. (2010). Learning-associated gamma-band phase-locking of action–outcome selective neurons in orbitofrontal cortex. *Journal of Neuroscience*, *30*(30), 10025–10038. https://doi.org/10.1523/JNEUROSCI.0222-10.2010

Witten, I. B., Steinberg, E. E., Lee, S. Y., Davidson, T. J., Zalocusky, K. A., Brodsky, M., … Deisseroth, K. (2011). Recombinase-driver rat lines: Tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*, *72*(5), 721–733. https://doi.org/10.1016/j.neuron.2011.10.028

Wong, S. A., Randolph, S. H., Ivan, V. E., & Gruber, A. J. (2017). Acute δ-9-tetrahydrocannabinol administration in female rats attenuates immediate responses following losses but not multi-trial reinforcement learning from wins. *Behavioural Brain Research*, *335*, 136–144. https://doi.org/10.1016/j.bbr.2017.08.009

Wong, S. A., Thapa, R., Badenhorst, C. A., Briggs, A. R., Sawada, J. A., & Gruber, A. J. (2017). Opposing effects of acute and chronic d-amphetamine on decision-making in rats. *Neuroscience*, *345*, 218–228. https://doi.org/10.1016/j.neuroscience.2016.04.021

Zeeb, F. D., Robbins, T. W., & Winstanley, C. A. (2009). Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. *Neuropsychopharmacology*, *34*(10), 2329–2343. https://doi.org/10.1038/npp.2009.62

Zhang, D., Yang, S., Yang, C., Jin, G., & Zhen, X. (2008). Estrogen regulates responses of dopamine neurons in the ventral tegmental area to cocaine. *Psychopharmacology*, *199*(4), 625–635. https://doi.org/10.1007/s00213-008-1188-6

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2921–2929. https://doi.org/10.1109/CVPR.2016.319