Sex-Differences in Cocaine-Induced Reinforcement Learning Deficits

Noemie M-L. Eustachon

Department of Psychology, McGill University

July 2024

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

Table of Contents

Abstract

Rationale: Chronic cocaine use is associated with poor cognitive flexibility and impairments in reinforcement learning (RL), with specific decrease in the sensitivity to positive and negative outcomes.

Objectives: The impact of sex on the development of cocaine-induced RL deficits remain poorly understood. Here, we aimed to first compare the learning strategies of male and female mice throughout the acquisition of a decision-making task and then characterize the emergence of cocaine-induced impairments in positive and negative RL in male and female mice.

Methods: Male and female mice were first trained to acquire a deterministic reversal learning task (DRL). The impact of daily cocaine injections (20 mg/kg) on mice's sensitivity to positive and negative reinforcers was then assessed using a probabilistic reversal learning (PRL) task. Drug or saline injections were administered after completion of the PRL task each day for 2 weeks.

Results: Female mice earned more rewards than males overall and displayed better integration of positive reinforcement on a DRL task. During the two weeks of injections and following two additional weeks of forced abstinence, cocaine decreased choice accuracy in both sexes. The sensitivity to negative outcomes, measured by the lose-shift ratio, transiently decreased during the first week of cocaine injections. Males only showed a lasting decrease in their sensitivity to positive outcomes emerging in the second week of injections, assessed by both the win-stay ratio and positive learning rate in a reinforcement learning modeling.

Conclusion: Chronic cocaine exposure impairs positive and negative reinforcement learning in a time and sex-dependent manner.

Résumé

Justification et Objectifs: La consommation chronique de cocaïne est associée à des déficiences dans l'apprentissage par renforcement, avec une diminution spécifique de la sensibilité aux conséquences positives et négatives. Nous avons cherché à comparer les stratégies d'apprentissage des souris mâles et femelles tout au long de l'acquisition d'une tâche de prise de décision, puis à caractériser l'émergence de déficiences induites par la cocaïne dans l'apprentissage par renforcement positif et négatif chez les souris mâles et femelles.

Méthodes : Des souris mâles et femelles ont été entraînées à acquérir une tâche déterministe d'apprentissage par inversion (DRL). L'impact des injections quotidiennes de cocaïne (20 mg/kg) sur la sensibilité des souris aux conséquences positives et négatives ont été ensuite été évalué à l'aide d'une tâche d'apprentissage probabiliste par inversion. Les injections de drogues ont été administrées après l'exécution de la tâche d'apprentissage par inversion probabiliste chaque jour pendant 2 semaines.

Résultats : Les souris femelles ont obtenu plus de récompenses que les mâles et ont montré une meilleure intégration du renforcement positif lors de l'acquisition du DRL. Pendant les deux semaines d'injections et après deux semaines supplémentaires d'abstinence forcée, la cocaïne a diminué la précision des choix chez les deux sexes. La sensibilité aux résultats négatifs a diminué transitoirement au cours de la première semaine d'injection de cocaïne. Les males ont ensuite démontré une diminution durable de leur sensibilité aux résultats positifs à partir de la deuxième semaine d'injection.

Conclusion : L'exposition chronique à la cocaïne altère l'apprentissage par renforcement positif et négatif en fonction du temps et du sexe.

Acknowledgements.

I am grateful for my supervisor, Jonathan Britt, for trusting me and supporting me throughout my years in the lab. He valued my opinions and encouraged me through every step of my experiments, always listened to any ideas I would bring, and showed care and trust for my future. I would also like to thank Milan Valyear for being the most supportive and attentive mentor; for being excited about my work when I was not, and for encouraging me to continue in science. I would also like to thank my friend and lab mates in Stewart Biology, Madeleine Morris, Scilia Sciandra, Heike Schuler and Serena Wu for bringing fun and warmth in a (crumbling) work building and always being there for me even when I was crying over broken Med levers. Finally, all experiments conducted in this lab were conducted with the help of a fantastic and driven team of undergraduate students, Allie Bernstein, Sanjna Puri, Frederique Marcoux, Willem Rosenberg and Rhea Xinxo.

Contribution of Authors.

The author of this thesis collected and analyzed the work presented throughout the entirety this paper with the guidance of Jonathan Britt. These contributions include designing and conducting behavioural experiments, breeding and weaning mice, coding for experiments, statistical analyses and preparation of the figures. Behavioural experiments were performed with the help of two undergraduate students in the lab, Allie Bernstein and Sanjna Puri.

1. Decision-Making Deficits in Cocaine Use Disorders

1.1 Substance Use Disorders: Prevalence and Existing Treatments.

Substance Use Disorders (SUDs) involve maladaptive substance use, leading to harm in physical and mental health and social relationships (DSM-5, APA, 2013; ICD-11, WHO, 2022). SUDs are a leading cause of global mortality; smoking accounted for 7.7 million premature deaths in 2019, alcohol use 2.4 million, and other substances 550,700 (Volkow & Blanco, 2023). Cocaine use disorder (CUD) is the fifth most prevalent SUD (Volkow & Blanco, 2023) with a 17-20.9% transition rate from use to abuse (Lopez et al., 2011) and 5.8 million cases in 2016 (Degenhardt et al., 2018). In 2022, 27,569 overdose deaths in the US were related to cocaine. Despite this, no pharmaceutical treatments are approved by the US Food and Drug Administration for CUD (Kampman, 2019)

Behavioral treatments, specifically cognitive behavioral therapy (CBT) and contingency management (CMT), have shown some efficacy in managing cocaine abuse. CBT focuses on preventing relapse and identifying situations associated with drug use, teaching patients coping strategies like recalling the negative consequences of drug use or substituting cravings with positive thoughts. Clinical trials support CBT for treating cocaine (Carroll et al., 1991; Maude-Griffin et al., 1998; Rohsenow et al., 2000). CMT promotes abstinence using positive reinforcement, providing vouchers redeemable for food and services upon reaching therapeutic goals. Vouchers compensate for the delayed benefits of treatment by offering short-term rewards (Kampman, 2019). Although CMT is effective in treating cocaine use disorder, it is costly and its effects fade once reinforcers are removed (Kampman, 2019).

Despite progress in developing psychosocial treatments, relapse remains a significant issue. Cocaine users following effective treatments have relapse rates of 60 to 70% at follow-up (Bisaga

et al., 2010; Dutra et al., 2008). Thus, there is a crucial need for better interventions for cocaine use disorder.

1.2 Decision-Making Impairments Associated with Cocaine Abuse.

Suboptimal decision-making is a key factor in the development and maintenance of SUDs. Impaired control over substance intake and the escalation of use despite awareness of its negative consequences are criteria for the clinical diagnosis of SUD, as defined by both DSM-5 (American Psychiatric Association, 2013) and the ICD-11 (WHO, 2022). Impaired control has a direct impact on patient's lives and will damage their social relationships, health, and ability to fulfill daily responsibilities.

1.2.1 Outcome Value Impairments

CUD patients not only overvalue substances over their daily life responsibilities but also display impairments in valuing natural reinforcers in the absence of drug rewards. Risk-taking behaviour and the valuation of immediate vs delayed rewards have been widely assessed in lab settings using the Iowa Gambling Task (IGT) (Bechara et al., 1994). Participants choose cards from four decks, each yielding monetary gains and losses. Two of the decks are disadvantageous in the long-term, yielding larger rewards but also large penalties, while the other two decks are advantageous with smaller rewards and penalties. The IGT score is calculated as the number of cards chosen from the advantageous decks, minus the ones from disadvantageous decks. A low IGT score is associated with impairment in sensitivity to future consequences; immediate large rewards are preferred over smaller long-term ones, and larger losses are devalued. Lower IGT scores have been reported in CUD patients compared to controls (Balconi et al., 2014; Barry & Petry, 2008; Kjome et al., 2010; Nigro et al., 2021; Verdejo-Garcia et al., 2007). This difference

persists after 25 days of abstinence and is positively correlated with heavier cocaine use (Verdejo-Garcia et al., 2007). Cocaine users thus display impaired sensitivity to future consequences of actions, favoring immediate rewards despite punishments even in non-drug-related settings. Impaired outcome valuations could affect patients daily lives and impact their ability to maintain abstinence and seek treatments; worse performance on the IGT is associated with both a higher rate of relapse up to 3 months later among cocaine users (Nejtek et al., 2013; Schmitz et al., 2009; Verdejo-Garcia et al., 2014) and early dropout from treatment (Stevens et al., 2013).

1.2.2 Behavioural Flexibility Impairments

While the IGT has highlighted impaired integration of the future consequences of actions and the valuation of outcomes among cocaine users, reversal learning paradigms have been used to measure behavioural flexibility deficits. Navigating daily life not only requires weighing outcomes, but also adapting to changes in the environment. In reversal learning tasks, subjects first acquire a cue-outcome association, such as the pairing of a visual stimulus with positive or negative feedback. Once a learning criterion is reached, contingencies are reversed, requiring subjects to adjust behavior and update existing cue-outcome associations. A main outcome measure is the number of incorrect choices following reversal, which is indicative of perseverative errors. CUD patients perform worse than controls on reversal learning tasks and show increased perseveration after reversal (Ersche et al., 2008, 2011; Fernández-Serrano et al., 2012; Verdejo-Garcia et al., 2015). The duration of cocaine use positively predicts these deficits (Verdejo-Garcia et al., 2015), with effects lasting after 15 days of abstinence (Fernández-Serrano et al., 2012). These difficulties in disengaging from previously learned behaviour have been conceptualized to reflect impulsive traits (Franken et al., 2008) and compulsion (Izquierdo & Jentsch, 2012), and they could explain the loss of control over substance intake.

1.2.3 Limitations of Traditional Cognitive Task Measures

While both the IGT and reversal learning tasks have advanced our understanding of the behavioral and cognitive changes that follow substance exposure, they are criticized for lacking construct validity and reliability. These cognitive measures, such as the IGT score and number of perseverative errors, are simultaneously influenced by different aspects of decision-making (Verdejo-Garcia et al., 2019). For instance, poor performance could reflect specific deficits in valuing and integrating rewards and/or losses. Choosing the disadvantageous deck in the IGT can be driven by hypersensitivity to gains, hyposensitivity to losses, or a combination of both. Furthermore, performance can also be influenced by different behavioral strategies; one can withhold response to a more advantageous option to explore different choices or, conversely, persevere with a previously reinforced option. While conventional cognitive measures highlight gross decision-making deficits, an in-depth characterization of impairments could help pave the way for the implementation of novel behavioral therapeutic approaches.

1.3 Reinforcement Learning Modeling.

Computational modeling provides a mechanistic description of behaviour by decomposing performance into different behavioural components. Each subprocess describes different aspects of a decision-making strategy, such as outcome sensitivity or choice consistency, and is defined by a set of parameters. Estimating parameter values and contrasting them across individuals enables the comparison of decision- making mechanisms between populations.

1.3.1 Reinforcement Learning: Formalism.

Reinforcement Learning (RL) is an increasingly popular decision-making model in which behaviour is optimized by maximizing gains and minimizing losses. An estimate of the expected value of each option is kept online. When presented with novel evidence, a reward prediction error (RPE) is computed as the difference between the outcome expected and the one observed. If an outcome is better than expected, the RPE will be positive, and the estimated value of the option will increase. Conversely, if an outcome is worse than expected, the RPE will be negative, and the estimated value will decrease. To update an existing estimated value, RPEs are added/subtracted to the existing value after being weighted by a learning rate parameter bounded by 0 and 1. A value close to 0 indicates inflexible learning; values are never updated despite being presented with conflicting evidence. Values closer to 1 represent immediate learning; in response to new evidence, previous estimates are directly updated. Adaptive decision-making requires a good balance of the learning rate value and depends on the environment. This process of updating estimated values is called value-updating. The value-updating process can use distinct learning rates for positive and negative outcomes, which are used as measures of the sensitivity to positive and negative outcomes, respectively.

RL models compute the value-updating process separately from action-selection. Actions are chosen based on their estimated values, and if the best option is chosen, the individual is displaying exploitative behaviour. Conversely, exploration refers to sampling from the environment. Exploration and exploitation can be balanced using a softmax policy, in which the inverse temperature parameter governs the ratio of exploration. A value close to 0 renders all actions equiprobable, and action choice is random, irrespective of value estimation. On the other

hand, a higher value of the inverse temperature corresponds to an exploitation strategy, where there is an increased probability of choosing higher valued actions.

1.3.2 Asymmetric Value Updating Impairments in Cocaine Use.

Fitting RL models to CUD patients revealed impaired learning from negative outcomes compared to controls. In probabilistic tasks, in which a stimulus is associated either with financial gains or losses, cocaine users performed worse than controls and showed diminished learning from loss, indicating decreased integration of punishments (Lim et al., 2021). This result aligns with studies showing decreased accuracy among cocaine users in avoiding an electric shock despite extensive training (Ersche et al., 2016) and in actively avoiding the disadvantageous deck on the IGT (Thompson et al., 2012). Withdrawal does not influence this effect; after cocaine users abstain for 72h, they show unchanged learning from large losses. (Wang et al., 2019). Collectively, these studies underscore a diminished sensitivity to financial and physical punishment among cocaine users. Yet, in an experimental design in which the negative outcome is the absence of financial gain, cocaine users had similar negative learning rates compared to controls (Lim et al., 2021), suggesting a distinction between punishment and an absence of reward, possibly arising from the saliency of the outcome.

RL modelling assesses both positive and negative learning rates in individuals. Cocaine users exhibit asymmetrical learning from positive and negative outcomes, with a slightly lesser reduction in learning rates for financial gains compared to losses (Lim et al., 2021). Interestingly, this reduction in learning from rewards is reversible; cocaine users asked to abstain for 72h showed an increase in positive learning rate (Wang et al., 2019). And while cocaine users initially showed decreased accuracy in identifying stimuli associated with positive feedback, extensive training was able to remediate that effect (Ersche et al., 2016).

Altogether, cocaine users display a decrease in sensitivity to positive and negative outcomes, but the decrease in positive reinforcement learning can be reversed with abstinence and further training. Based on the resistance of cocaine users to negative reinforcement, therapeutic approaches need to focus on positive reinforcement to maintain abstinence. This effect reinforces the theory behind contingency management, which focuses on maintaining abstinence with food and service vouchers and has proven effective for treating CUD (Kampman, 2019).

1.4 Limitations of clinical work and Outstanding Questions.

1.4.1 Causality of Substance Use on RL Impairments

Clinical studies have reported clear decision-making deficits among cocaine users compared to controls, with a stark decrease in the sensitivity to punishment and a smaller reduction in sensitivity to positive reinforcers. However, several limitations are to be noted. Due to the design of these studies, it is unclear whether such impairments are a *consequence* of CUD or a *predisposition* to developing one and whether these impairments progress with ongoing cocaine use. No differences in reinforcement learning parameters were reported between controls and patients with a more recent cocaine use disorder diagnosis (i.e., less than 4 years on average (Zühlsdorff et al., 2024) versus 17.1 years (Wang et al., 2019) and 13.7 years (Lim et al., 2021) highlighting that decision-making deficits may arise with longer chronic cocaine exposure. Moreover, abstaining from drug use for several months (average of 2.73 months) renormalizes reinforcement learning rate parameters, unlike a 72h withdrawal period (Lim et al., 2021) . This pattern of results is reinforced by the correlation between duration of cocaine use and the strength of cognitive deficits on the IGT (Verdejo-Garcia et al., 2015).

1.4.2 SUD comorbidities and RL impairments

SUDs frequently co-occur, complicating the analysis of their individual effect on decision-making. Research indicates that alcohol, nicotine, methamphetamine, marijuana, heroin, and polysubstance users all present deficits on the IGT (Gowin et al., 2018), and subtle differences in RL parameters have been shown. Alcohol users display decreased positive learning rates and enhanced learning from negative outcomes (Bağci et al., 2022), in contrast with the dampened sensitivity to punishment among cocaine users (Ersche et al., 2016; Lim et al., 2021; Thompson et al., 2012). Patients with a past history of opioid use show no difference in learning rates or exploration, but a general tendency to repeat previous actions (Myers et al., 2016). Smokers however present with a similar phenotype as cocaine users. Lower learning rates for negative outcomes was found in current smokers compared to controls (Rai et al., 2019) and in smokers relapsing after abstinence (Baker et al., 2018). Additionally, decreased learning rates for positive outcomes were present in current and ex-smokers (Rai et al., 2019) as well as in abstaining smokers (Baker et al., 2018). Finally a study using a sample of mixed substance users reported that while substance users in general displayed lower learning rate from negative outcomes, stimulant users in particular had higher positive learning rates and greater exploration compared to individuals with other substance use diagnoses (Taylor et al., 2023). Overall, these studies point to impaired value-based decision-making in substance use disorders, with specific impairments varying with the substance abused. Stimulant users in particular may display different RL profile from other substance users (Bağci et al., 2022; Taylor et al., 2023). These findings may hint to substance-specific effects on behaviour. While substances share common mechanisms, their precise effects vary at both the behavioural and neural levels (Volkow et al., 2019). Alternatively, it is possible

that these different RL deficits are not a consequence of SUDs, but rather a predisposition. Different RL profiles at baseline might draw individuals to abusing different substances.

While all aforementioned studies on the impact of cocaine abuse on decision-making only included participants with cocaine as their main abused substances, most included individuals actively using cannabis, alcohol or tobacco. Due to the difficulties in finding samples of 'pure' users, the effect of cocaine uses, and other comorbid disorders cannot be disentangled. Due to the variety of RL impairments across substances, it is crucial to characterize substance effects to develop treatments adapted to individuals.

SUD patients also have a high rate of comorbidities with anxiety disorders and depression (Bethesda, 2020), which are also associated with changes in reinforcement learning (Brown et al., 2021). While clinical samples have the advantage of being more representative of the general population living with a CUD, they are confounded by co-diagnoses and differential patterns of substance use. Sample heterogeneity can be accounted for *within* studies, but drawing conclusions *across* studies is challenging. To answer questions regarding the evolution of decision-making deficits with cocaine intake, we hence turn towards pre-clinical models of SUDs.

2. Chronic Cocaine Exposure and Decision-Making: Insights from Animal Models

2.1 Modeling SUD and Assessing Decision-Making in Rodents.

Rodent models of cocaine use disorders involve exposing animals to the drug through self-administration, where they press a lever to self-infuse cocaine, or via experimenter-administered injections. Decision-making is assessed using simplified, rodent-adapted versions of human cognitive tasks. Instead of monetary gains or losses, palatable food serves as a positive reinforcer, and bitter food or electric shocks serve as punishment.

2.2. Reinforcement Learning Impairments as a Consequence of Chronic Cocaine Exposure.

Chronic cocaine exposure has been shown to cause decision-making deficits in rodents. After seven to fourteen days of cocaine treatment and up to four weeks of abstinence, both mice and rats displayed impaired performance on reversal learning paradigms and perseverative responses to devalued outcomes, whether cocaine was self-administered (Bechard et al., 2018; Calu et al., 2007) or injected (Himanshu Gangal et al., 2023; Krueger et al., 2009; LeBlanc et al., 2013; McCracken & Grace, 2013; Schoenbaum et al., 2004; Schoenbaum & Setlow, 2005). Interestingly, individual differences were reported in the rodent IGT, as only a subset of rats showed decreased choice of the advantageous option after 19 days of self-administration. Impaired and non-impaired animals had similar cocaine intake and escalation rate, but impaired rats were more susceptible to relapse (Cocker et al., 2020). This evidence suggests a direct causal effect of chronic cocaine exposure on decision-making, with deficits predicting relapse susceptibility. Consistent findings across exposure protocols indicate that all patterns of cocaine use impact decision-making, with individual differences in the severity of these deficits.

RL models can be used to analyze rodent behavior and extract parameter estimates. Studies report that rats with low and high escalation of cocaine self-administration over six days had different learning strategies eight days later. High-escalation rats were less accurate in adjusting behavior after negative reinforcement compared to both the low escalation group and controls. Although RL modeling showed no difference in sensitivity to reinforcers, high-escalation rats failed to exploit learned options and tended to repeat actions (Zhukovsky et al., 2019). After 21 days of self-administration, rats similarly showed decreased behavioral adjustment following negative reinforcement, which was reflected in decreased sensitivity to negative reinforcement in RL models (Groman et al., 2020a). Rodent models highlight a decreased ability to learn from

negative reinforcement due to substance exposure, mirroring the reduced punishment learning seen in clinical samples. Cocaine users only showed decreased learning from financial loss or electric shocks, but not from the absence of financial outcomes (Ersche et al., 2016; Lim et al., 2021; Thompson et al., 2012). In contrast, rats showed impaired adjustment following the absence of a food reward (Groman et al., 2020a, Zhukovsky et al., 2019). Since rats are food-restricted, the absence of food reward may be more salient than the absence of a loss/win for humans.

2.3. Reinforcement Learning Impairments as a Predisposition to Chronic Cocaine Exposure.

The predisposing effects of suboptimal decision-making on future substance use are measured in rodents by assessing decision-making prior to administering cocaine. Studies have shown that only the rats that prefer riskier options on a gambling task at baseline showed further performance deterioration following cocaine self-administration as well as greater incubation of craving after 30 days of withdrawal (Ferland & Winstanley, 2017). Similarly poor performance on a PRL task predicted greater future escalation in cocaine-self-administration in rats (Groman et al., 2020b). Thus, suboptimal decision-making at baseline may predispose one to developing an SUD later on.

RL modeling studies on patients with cocaine use disorders show a stark decrease in sensitivity to punishment and a marginal reversible decrease in positive reinforcer sensitivity. Modeling of rats' behavior on a PRL task at baseline showed that a lower sensitivity to positive but not negative reinforcement was negatively correlated with future escalation of use, such that a lower sensitivity to rewards predicted greater escalation of cocaine self-administration (Groman et al., 2020b). The same laboratory trained a cohort of rats to perform a PRL task during adolescence and compared their learning trajectory with cocaine self-administration in adulthood.

Rats that showed a smaller increase in sensitivity to positive reinforcement in adolescence displayed greater escalation in adulthood, highlighting a causal link between the development of a reinforcement learning strategy and cocaine abuse (Villiamma et al., 2022). Preclinical research parallels clinical findings that show individuals with cocaine use disorders had reduced learning from positive non-drug reinforcers, suggesting this deficit may predispose them to developing an SUD.

2.4. Outstanding Questions.

Preclinical research has thus revealed a causal link between impaired positive reinforcement learning and future escalation in cocaine use as well as between cocaine exposure and the later development of negative reinforcement learning impairments. However, several questions remain. First, the presented work has not characterized the development of impairments throughout chronic cocaine exposure, only prior to or after. CUD patients are faced with choices in their daily life, and concurrent drug use while navigating complex environments may exacerbate the impact of drug use on decision-making. It is also possible that impairments in decision making evolve throughout cocaine exposure. Additionally, none of the aforementioned work has explored sex as a meaningful variable with respect to the effect of cocaine on RL, despite sex and gender differences in the presentation of SUDs and in healthy decision-making.

3. Sex-Differences in SUDs and Decision-Making.

3.1 Gender and Sex Differences in SUDs.

SUDs are more prevalent among men than women worldwide, with an estimated 2.3 to 1.5 higher ratio of SUD in men compared to women (Degenhardt et al., 2018). Out of the 5.8 million cases of cocaine use disorders in 2016 worldwide, 68% were men (Degenhardt et al., 2018).

However, this gender gap has been narrowing, and countries with lower traditional gender role are associated with a smaller gender gap in SUD prevalence (Seedat et al., 2009). Women transition from first substance use to abuse faster than men (Becker et al., 2017). A similar effect of sex has been reported in rodents; females showed faster escalation and greater intake of cocaine than males, and they required lower doses to develop a preference for a spatial location associated with cocaine exposure (Becker & Koob, 2016).Women are also more likely to relapse than men, have shorter periods of abstinence, display greater withdraw symptoms with parallel findings in rodents (Becker et al., 2017). Despite a clear impact of sex and gender on the trajectory of SUDs, both clinical and preclinical studies have extensively focused on males.

3.2 Gender and Sex Differences in RL.

Gender and sex-differences have been reported in baseline decision-making. Men choose more cards from the advantageous deck than women on the first 100 trials of the IGT, but women reach the same level after an additional 40-60 trials on average (Van Den Bos et al., 2013). A similar pattern was reported in rats tested on the rodent IGT (Van Den Bos et al., 2012). This, however, does not indicate faster learning overall in males. In fact, female mice have been shown to acquire a probabilistic decision-making task faster than males and display alternative learning trajectories over time (Chen, Ebitz, et al., 2021). Furthermore, despite showing similar accuracy once the task is acquired, males and females employ different learning strategies, with males exhibiting greater exploration and reduced outcome integration Chen, Knep, et al., 2021). The evidence suggests not better learning overall in a specific sex, but rather sex differences in behavioral and learning strategies.

Research has shown sex-differences in animals' sensitivity to positive and negative reinforcement. While female mice display similar speed of acquisition of responding for sucrose reinforcers as males, they are slower learn to avoid a foot shock (Kutlu et al., 2020). This research indicates there is lower negative reinforcement learning rates among females, but contrasting evidence has been demonstrated when animals have to weigh rewards over punishment. In a task in which animals have the choice between a food reinforcer presented concurrently with a foot shock (dubbed as a 'riksy' option) and a safer but smaller reinforcer, females display increased risk aversion. Female mice and rats are less likely than males to choose a reward associated with a risk of footshock (Islas-Preciado et al., 2020; Kutlu et al., 2020; Liley et al., 2019; Orsini et al., 2016)

3.3 Sex Differences in SUD & Decision-Making.

The impact of sex on cocaine-induced decision-making is an understudied area, despite clear sex differences in cocaine administration behaviour and baseline decision-making. Among current cocaine users, women displayed greater deficits on the IGT than men (Van Der Plas et al., 2009). Abstinence may, however, similarly impact males and females, as 14 days of abstinence from cocaine self-administration similarly increased risky decision-making in male and female rats. (Blaes et al., 2022). Individuals differences in decision-making at baseline differentially impacted future cocaine intake across sexes; risk aversion predicted future increase in cocaine self-administration in females but not male rats (Orsini et al., 2020). Altogether, the current evidence points to potential impact of sex on mediating cocaine-induced decision-making deficits, but too little work has been conducted to draw clear conclusions. Furthermore, current work has primarily focused on sex-differences in risk-taking behaviour in tasks where reward is weighed against the possibility of punishment, but not in tasks where the negative outcome is an absence of reward.

4. Aims & Hypotheses

Preventing relapse remains an outstanding issue in the treatment of cocaine use disorders. Chronic exposure to cocaine has been associated with impairments in flexible decision-making when working for natural reinforcers in both humans and animals (Bechard et al., 2018; Ersche et al., 2008) with specific deficits in positive and negative reinforcement learning (Groman et al., 2020a; Groman et al., 2020b). However, no work until now has looked at how impairments can change with training throughout cocaine exposure. Cocaine users are constantly faced with everyday choices while using substances, and their reinforcement learning strategy may change with ongoing drug exposure during task training. Another underexplored avenue of research is the impact of biological sex on cocaine-induced reinforcement learning deficits. Males and females use different decision-making strategy in baseline conditions (Kutlu et al., 2020), and behaviour is differentially impacted by cocaine exposure (Van Der Plas et al., 2009). Sex is hence a potential mediator for the impact of cocaine on decision- making.

The experiments presented here aim to 1) characterize sex-differences in learning trajectories on a reversal learning paradigm and 2) identify the impact of chronic cocaine exposure on reinforcement learning in a sex and time dependent manner. To this end, wild-type C57BL/6J male and female were trained to perform a probabilistic reversal learning (PRL) task. Performances during task training are contrasted across sexes to identify sex-differences in learning. Following acquisition of the PRL task, mice were injected with 20mg/kg cocaine or saline for 14 consecutive days. Testing on the PRL was maintained throughout cocaine exposure and occurred 1h prior to the daily drug injections. Mice were then subjected to 14 days of forced abstinence from cocaine and tested for an additional week on the PRL task to identify potential long-term effect of cocaine on decision-making.

Methods

Animals.

All animals used in the present study were C57BL/6 wild-type mice (N=28, F=11, M=17, 6-8 weeks of age), group housed in standard laboratory conditions with a reversed 12-hour light/dark cycle and an ad libitum diet. All experiments were conducted in accordance with the Canadian Council of Animal Care and the McGill Animal Care Committee.

Apparatus.

Mice were trained in sound attenuating chambers that had levers available on either side of a centrally located food receptacle (Med Associates). A houselight was located on the opposite side of the chamber. All behavioral data were collected using Med Associates software.

Pre-Training.

The training for the behavioral task was divided into two phases: habituation and fixed ratio (FR) schedule of reinforcement. During habituation, all mice were first exposed to 10% sweet-condensed milk (SCM) in their home cage for three hours, one hour into their dark-light cycle, for three consecutive days. Mice were then habituated to the testing chambers for one 35-minute session while a 10% SCM solution was freely available in the food receptacle. Licking in the food receptacle was paired with illumination of the food port. In the next phase of training, mice were again placed in the operant box and reward delivery was contingent on lever pressing on either of the two lever according to a fixed ratio (FR1) schedule of reinforcement. Both levers extended at the start of the trial and retracted for 3 seconds after a lever press or after 60s elapsed. Delivery of SCM was paired with illumination of the central food receptacle. All mice were trained for 2h hour daily until they made a lever press on at least 40% of trials within a session. Subsequent sessions were 35 minutes long, with an engagement criterion set to answering at least 65% of trials.

Deterministic Reversal Learning Task.

Training sessions on the deterministic reversal learning (DRL) task were limited to one hour. Both levers were extended for 20 seconds maximum with a 3s ITI. The reinforcement probabilities of the two levers were set to 100% and 0%. A lever press on the non-reinforced lever was punished by the illumination of the house light for 3 seconds. Correct lever presses were reinforced by the illumination of the food port for 3s along with the delivery of 10% SCM. Lever reinforcement probabilities reversed after 5 consecutive presses on the reinforced lever.

Probabilistic Reversal Learning Task

All mice were tested for 1h daily on a Probabilistic Reversal Learning (PRL) task. Two levers on each side of the food port extended at a start of each trial. The levers retracted for 3 seconds following a press of either lever or after 20 seconds elapsed. The two levers had a probability of reinforcement of 80% and 20%. These probabilities reverse after 5 consecutive presses on the high probability lever. Reinforced lever presses triggered delivery 10% SCM and illumination of the food receptacle. Non-reinforced lever presses were punished by illumination of the house light for 3 seconds.

Injection Protocol

All mice were injected with saline at the end of each week of initial training to habituate them to injections. During the experiment, mice were administered with an intraperitoneal injection of either saline (N=10, F=6) or 20mg/kg cocaine hydrochloride (N=10, F=5) in a 10 ml/kg injection volume dissolved in saline. Injections occurred daily for 14 days, 1 hour after testing on the PRL. Locomotor activity was recorded for 25 minutes post-injections.

Tasks Variables.

Task variables included number of reversals completed, number of rewards earned, number of lever presses, reaction times (defined as time elapsed since the start of the trial and the first lever press), and reward collection latency (defined as time elapsed between lever press and first head-entry in the central food port). Additionally, side bias was computed as $\frac{|Press_{left} - Press_{right}|}{Press_{left+right}}$. Win-stay and lose shift ratios were defined as the average probability of pressing the same lever twice given the previous trial outcome.

Statistical Analyses of Sex-Differences in DRL.

All statistical analyses were conducted through custom python and R scripts. Data visualization and plotting were performed using the seaborn library in python (Waskom, 2021). To assess sex differences in learning over sessions, linear mixed-effects models were fit on session-average data in R using the lme4 package (Bates et al., 2015), optimizing the restricted maximum likelihood. Equation 1 describes the model used to model variable *Y* of mouse *i* on session *j* with a random intercept μ for each mouse *i* and a residual error $\epsilon_{i,j}$.

$$Y_{i,j} = \beta_0 + \beta_1 Sex_i + \beta_2 Session_j + \beta_3 (Sex_i \; x \; Session_j) + \mu_i + \epsilon_{i,j} \qquad (Equation\ 1)$$

β_0, β_1, β_2, and β_3 correspond, respectively, to the intercept, the differences in *Y* of males relative to females across sessions, the slope of change of *Y* over sessions across sexes, and the interaction of sex and session. The significance of coefficients was assessed using t-tests computed with Satterthwaite's method.

For trial-by-trial analyses of performance following reversals, the 4 trial sequences following reversals were extracted, ensuring they did not overlap with the 4 trials preceding reversals. Generalised-linear mixed effect models (GLMMs) were fit by maximum likelihood independently on each training session to predict reward (R) as a function of the number of trials

elapsed since reversal and sex, with random effects of $u_k^{(1)}$ reversal and $u_i^{(2)}$ mouse. Equation 2 describes model formulation to compute the probability of earning a reward *R* at trial *t* for reversal *k* and mouse *i*.

$$P(R_{itk} = 1) = \frac{1}{1 + \exp\left(-\left(\beta_0 + \beta_1 \text{Sex}_i + \beta_2 \text{trials}_t + \beta_3 (\text{Sex}_i \times \text{trials}_t) + u_k^{(1)} + u_i^{(2)}\right)\right)}$$

$$(Equation\ 2)$$

Statistical Analyses of Cocaine Effects on PRL.

All statistical analyses were conducted in custom python and R scripts. Sessions were binned into 4 time points: Baseline (the first seven sessions), Week 1 (the first seven sessions following drug injections), Week 2 (the next seven sessions following drug injections), and Post-Abstinence (the 5 sessions following abstinence). Linear mixed-effect models were used to compare drug group, sex, time as well as their three-way interactions on task variables, with random effects of sessions and mouse. The behaviour of males and females were only graphed separately if a significant interaction with sex was detected. The data was otherwise combined to improve the statistical power of the analyses.

Reinforcement Learning Modelling.

The choice behaviour of mice in the PRL task across sessions was concatenated across time points and modelled with a double Q-learning model with forgetting (DQF). This model fit better than other reinforcement learning models, with lower Akaike Criterion's Criteria (see Supplemental Material, Table 1 and Figure 1 and 2). Value-updating of each lever *L* at trial *t* is

described in Equation 3. Three learning rates were used: a positive learning rate $\propto^{+}$, a negative learning rate $\propto^{-}$, and a forgetting rate $\propto^{f}$, as described in equation 3.

$$Q_{t+1}^{L} = \begin{cases} Q_t^L + \alpha^{+}(r_t - Q_t^L) & if \quad a_t = L, r = 1 \\ Q_t^L + \alpha^{-}(r_t - Q_t^L) & if \quad a_t = L, r = 0 \\ Q_t^L + \alpha^{f}(-Q_t^L) & if \quad a_t \neq L \end{cases} \qquad \text{(Equation 3)}$$

r_t corresponds to the reward earned at trial *t*, a_t corresponds to the action chosen at trial *t*. Actions were determined according to a softmax policy (Equation 4). The inverse temperature β controls the degree of exploration.

$$p(a_{t+1} = L) = \frac{e^{\beta Q_t^L}}{\sum e^{\beta Q_t^l}} \qquad \text{(Equation 4)}$$

Parameters were fit to the data via maximisation of the log likelihood.

Chapter 1: Sex-Difference in Learning.

Background.

Value-based decision making requires estimating the subjective values of available options and taking decisions accordingly. To navigate complex and changing environments, humans and animals alike need to identify relevant cues and update cue-outcome associations by integrating new evidence. This type of flexible decision-making is necessary to adapt to changes in cue-outcome contingencies. Impaired decision-making has been associated with many psychiatric disorders such as substance use disorders, major-depressive disorders, and schizophrenia (Caceda et al., 2014). Despite clinical and pre-clinical work having demonstrated sex-difference in the expression of these pathologies (Bangasser et al., 2021; Becker et al., 2017), females are still underrepresented in the study of flexible decision-making, so results and conclusions may only be applicable to males.

Rodent research has reported sex-differences in learning sensitivity from positive and negative outcomes. Females have been reported to be more motivated than males to self-administer both drug and non-drug rewards (Becker, 2008; Hu et al., 2004; Kutlu et al., 2020). Conversely, females are slower to learn to avoid shocks (Kutlu et al., 2020). While this indicates that females have enhanced and decreased sensitivity to positive and negative reinforcement, respectively, females have been demonstrated to weigh outcomes differently when faced with conflict. Female mice and rats are less likely than males to choose a larger reward associated with risk of foot shock over a safer, smaller reward (Islas-Preciado et al., 2020; Kutlu et al., 2020; Liley et al., 2019; Orsini et al., 2016), hence displaying increased risk-avoidance behaviour. Altogether, sex-differences in decision-making are a complex interaction of available choices, with greater sensitivity to reward and lower sensitivity to punishment when presented individually, but a preference for smaller, safer rewards over larger, riskier rewards when punishment and reward are presented concurrently.

These interactions between punishment and rewards raise the question of how males and females differ in decision-making in more 'naturalistic' settings, in which the punishment is not physical. In tasks where mice learn to associate visual cues with probabilistic delivery of rewards and wrong choices are punished by the absence of reward, female mice display faster task acquisition and are more likely to employ a systematic strategy over time. In contrast, males switch choices more often and are more likely to base their choices on the previous outcome. Females further display less explorative choices, faster integration of outcomes, and more adaptive adjustment of behaviour following previous trial outcomes (Chen, Ebitz, et al., 2021; Chen, Knep, et al., 2021)

Aims.

Prior work points to clear sex-differences in learning strategies in spite the absence of physical punishment. Here, we aim to investigate how male and female mice adapt their behaviour to changing contingencies in a reversal learning paradigm, in which the location of a reinforced lever reverses over time. Additionally, males and females in the aforementioned studies were food-restricted to increase motivation and task engagement, which may make the absence of reward more akin to a physical punishment. Further, these prior results might be biased by the confound that females are more willing than males to decrease their daily water and food intake. In an experimental setting where rats were housed in an environment where their nest was safe but resources were only available by traveling through an area that randomly delivered shocks, female but not male rats decreased their meal size and arrested their weight gains (Pellman et al, 2017), which is indicative of different valuation of metabolic needs over punishments.

In this current experiment, ad libitum fed male and female C57BL6/J mice were trained to perform a deterministic reversal learning task (DRL) for 10 sessions. By comparing the rate of

task acquisition and sensitivity to trial outcomes across sexes we demonstrate that females have faster learning rates and better performance following reversals.

Results

1. Males and Females Acquire an Appetitive Response at a Similar Rate.

Ad libitum fed wild-type male and female mice (N=27, F=11) were trained on a fixed ratio of reinforcement (FR1) schedule of reinforcement to acquire lever pressing behaviour. Trials started with the extension of two identical levers, which when pressed triggered the delivery of 10% sweet-condensed milk in the central food port, along with illumination of a cue light. The first three sessions were 2h long, and the following nine sessions lasted 30 minutes.

Lever pressing behaviour was acquired at a similar rate in both sexes during both the 2h (**Figure 1.A**) and 30 minute sessions (**Figure 1.B**) (linear mixed-effect models, 2h sessions: main effect of Session: β=44.071 $p<0.001$; Sex: β=19.3843, p=0.598; Sex x Session interaction: β=-12.024, p=30.784; 30 minutes sessions: main effect of Session: β=4.365, $p<0.001$; Sex: β=3.547, p=0.657; Sex x Session interaction: β=-1.509, p=0.0.90), with no sex-differences in side bias (**Supplementary Figure 3**). While the reaction time, defined as the latency between trial start and lever press, did not significantly differ between the sexes (**Supplementary Figure 3)**, females displayed a faster decrease in their latencies to retrieve rewards following a lever press than males

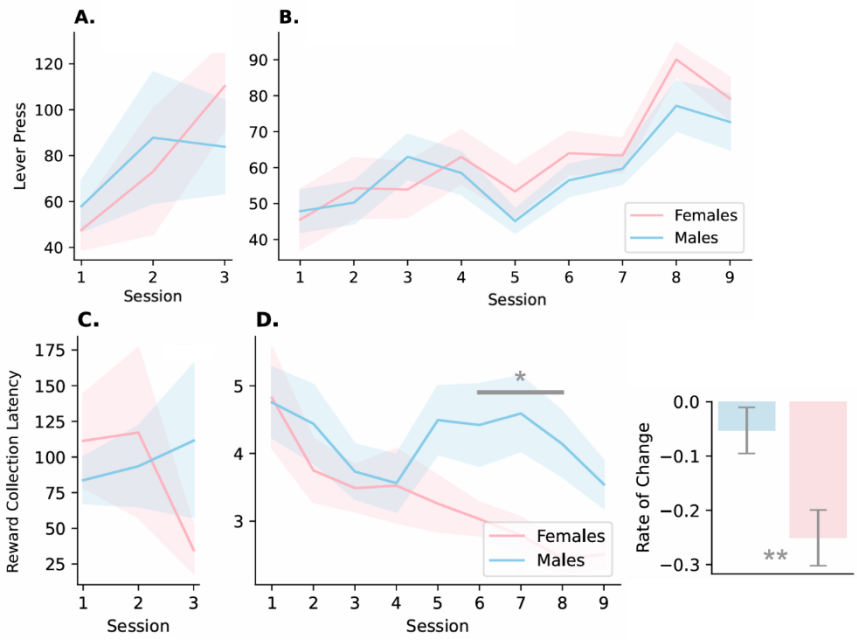


Figure 1. Training Sessions of FR1. Lever presses similarly increased in both sexes over **A,** 2h long sessions ($p<0.001$), and **B,** 30 minutes sessions ($p<0.001$). No sex-differences were detected in either session types**. C,** The latency to retrieve rewards did not change over time or vary as a function of sex during the 2h session. **D,** Females showed a faster decrease in their latency to retrieve rewards over the 30min training sessions. ($p=0.003$)

during the 30 minutes sessions (30 minutes sessions: Session: $\beta=-0.15177$, $p<0.001$; Sex: $\beta=0.0475$, $p=0.875$; Sex x Session: $\beta=-0.099$, $p=0.003$; post-hoc, females vs males $t(2.966)=0.198$, $p=0.003$). As reward delivery in the central food port is indicated by a 3s light cue immediately following a lever press, a faster latency to retrieve rewards among females may indicate higher certainty about visual cue-outcome association.

2. Females Perform Better on a Reversal Learning Task.

Following acquisition of lever pressing behaviour, all mice were trained on a reversal learning (RL) task for 10 sessions of 1 hour. During each session, one lever was randomly selected to be inactive. Lever presses on the inactive lever were punished by illumination of the house light for 3 seconds. The reinforced and non-reinforced lever locations reversed following 5 consecutive presses on the active lever (**Figure 2.A**.) The location of the active lever was not cued in any way so the most relevant consideration for the decision of which lever to press was the outcome of previous trials.

All mice learn to perform the task during training, with a similar increase between the sexes in both the number of reversals performed and number of rewards earned per session over time. While both males and females displayed similar learning rates, females performed better overall than males (**Figure 2.B.**; Reversals: fixed effect of Session, $\beta=1.106$, $p<0.001$; Sex: $\beta=2.749$, $p=0.099$;

no Sex x Session interaction β=-0.143, p=0.141; **Figure 2.C**; Rewards: fixed effect of Session, β=6.650 p<0.001, Sex: β=11.624, p=0.028; no Sex x Session interaction: β=-0.030, p=0.954). Higher rates of reward were not associated with increased engagement in the task, as rates of lever pressing similarly increased across sessions in both sexes (**Figure 2.D;** main effect of Session, β=4.853, p<0.001; no main effect of Sex, β=7.437. p=0.631, nor interaction: β=-0.768. p=0.458).

We next investigated whether males earned less rewards because they adapted poorly to contingency reversals. To analyze trial-by-trial changes in behaviour following reversals, generalized linear mixed effect models (GLMMs) were fit to predict the probability of earning rewards as a function of the number of trials passed following reversals as a function of sex (see Methods). During the first session of DRL, males and females displayed different performance

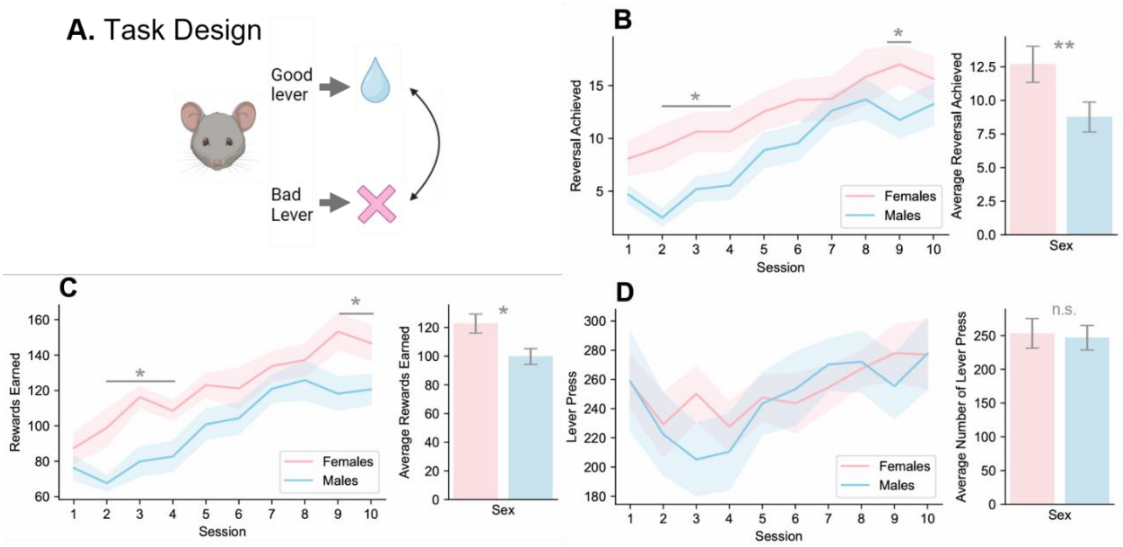


Figure 2. Task performance and engagement over sessions. A. Design of the deterministic reversal learning task. **B.** The number of reversals increased over training (p<0.001), with females achieving more reversals on average than males (p=0.001). **C.** The number of rewards earned increased over training (p<0.001), with females earning more rewards on average than males (p=0.028). **D.** The number of lever presses increased over sessions (p=0.011), but no sex-differences were detected (p=0.631)

trajectories following reversals. While females slightly increased their probability of earning rewards over trials, males showed the opposite patterns and performed more poorly over trials (**Figure 3.A and B,** GLMM, Trials, β=0.219, p=0.0492; Sex, β=1.132. p<0.001; Sex x Trial interaction, β=-0.3861. p=0.021, *f*emales vs males coefficients, z =2.299=p=0.021). However, by session two, both males and females showed similar learning curves following reversals, with increases in the probability of earning rewards over trials (**Figure 3.A,** Trials, β=0.350, p=0.001; Sex, β=0.533. p=0.233, Sex x Trials, β=-0.138, p=0.474). Mice continued to show increases in the probability of earning rewards over trials over the next few sessions, with no effect of sex. However, females showed steeper slope of adjustment by the last three sessions. with significantly faster adaptation to reversals than males (**Figure 3.A and C,** Session 10: Trials, β=0.767, p<0.001; Sex, β=0.187. p=0.550; Sex x Trials, β=-0.247, p=0.030, females vs males coefficients, z =2.1659, p=0.030). Therefore, males are initially worse than females to adapt to reversals but learn to adjust their behaviour with training. Females, however, display faster adjustment to contingency reversals over the last sessions.

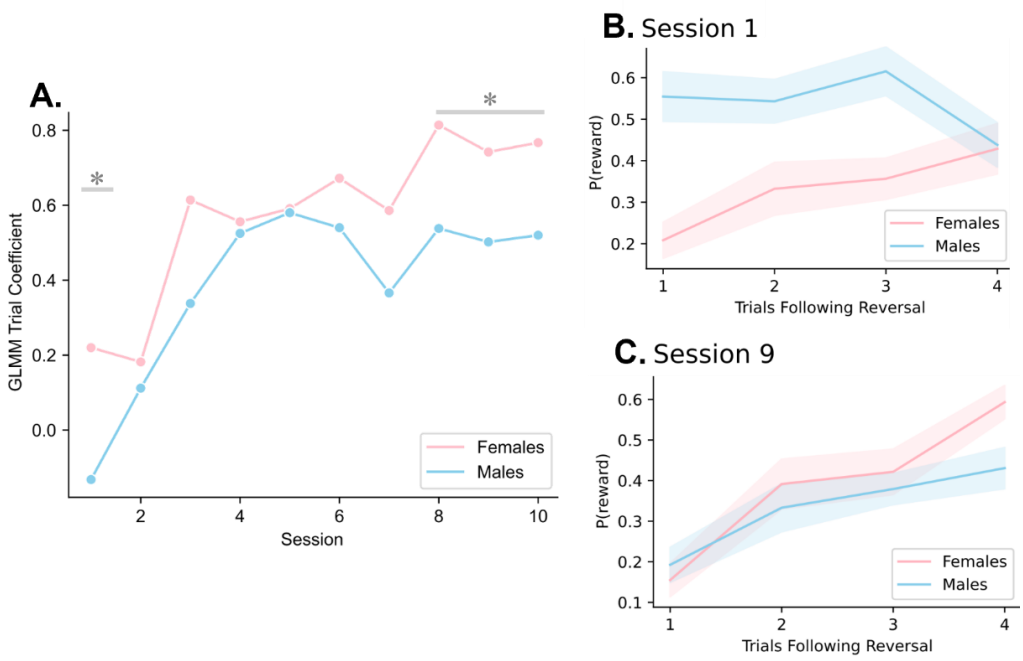


Figure 3. Sex-differences in Speed of Performance Adjustment Following Reversals. GLMMs were fit on each training session independently of whether the mouse earned a reward at that trial, as a function of sex and the number of trials following reversals . **A.** GLMM coefficient for the trial number, representing the slope of learning, in males and females across sessions. Males displayed slower learning rates during sessions 1, 8, 9 and 10. **B**. During the first session, males and females differentially adjust their behaviour ($p=0.02$). Males were less likely to earn a reward as trials increased, whereas females were increasingly more likely to earn a reward. **C.** In the final session of DRL, both sexes displayed an increased in performance following reversals, but there was a steeper slope among females ($p<0.004$).

3. Sex-Differences in Outcome Modulation of Behaviour.

Females displayed better adjustment to changes in task contingencies than males. We next investigated whether an underlying reason might be differential sensitivities to previous trial outcomes between sexes. Adaptive navigation of the RL task requires choice perseveration when the chosen lever is reinforced and a shift to the alternative lever when the chosen favored is not reinforced, which corresponds to a lower probability of pressing the same lever following lack of reinforcement (Lose-Stay) and a higher probability of pressing the same lever following reinforced trials (Win-Stay). Within the first three sessions, females had an initial higher win-stay ratio than males. While win-stay behaviour increased over sessions in both sexes, males displayed a faster rate of increase (**Figure 4A;** linear mixed effect model; fixed effect of Session, $\beta=0.014$, $p<0.001$; Sex, $\beta=0.048$. $p=0.013$; Sex x Session interaction: $\beta=-0.0037$, $p=0.024$; post-hoc test, females vs males, $t(241) =-2.277)=p=0.024$). Females were additionally faster to retrieve rewards overall following a reinforced lever press (**Figure 4D**; Sex: $\beta=-1.788$, $p=0.004$; Session: $\beta=-0.289$,

p<0.001; no Sex x Session interaction, β=0.070. p=0.281) and were faster to press a lever following a reinforced trial (**Figure 4C**; Sex: β=-0.545, p=0.012; Session: β=-0.179 p<0.001, no Sex x Session interaction: β=0.023, p=0.245). Overall, female mice were better at adjusting their behaviour following positive reinforcement than males. They displayed more correct choices following reinforced trials, faster reaction times to the visual cue indicating reward delivery, and more adaptive integration of previous rewards when making subsequent choices.

Following lack of reinforcement, females were able to accurately learn to shift responding to the other lever across sessions, while males continued to display perseverative behaviour. Lose-Stay ratios decreased over time in females but increased in males (**Figure 4B;** Sex: β=0.011, p=0.484; Session: β=-0.003, p=0.169; Sex x Session: β=-0.004 p=0.017; post-hoc, females vs males coefficients, t(241) =-2.404)=p=0.017). The reaction time following non-reinforced trials was similar in both sexes and did not change over time (Sex: β=0.098 p=0.635; Session: β=-0.045, p=0.0733; Sex x Session: β=0.022 p=0.382).

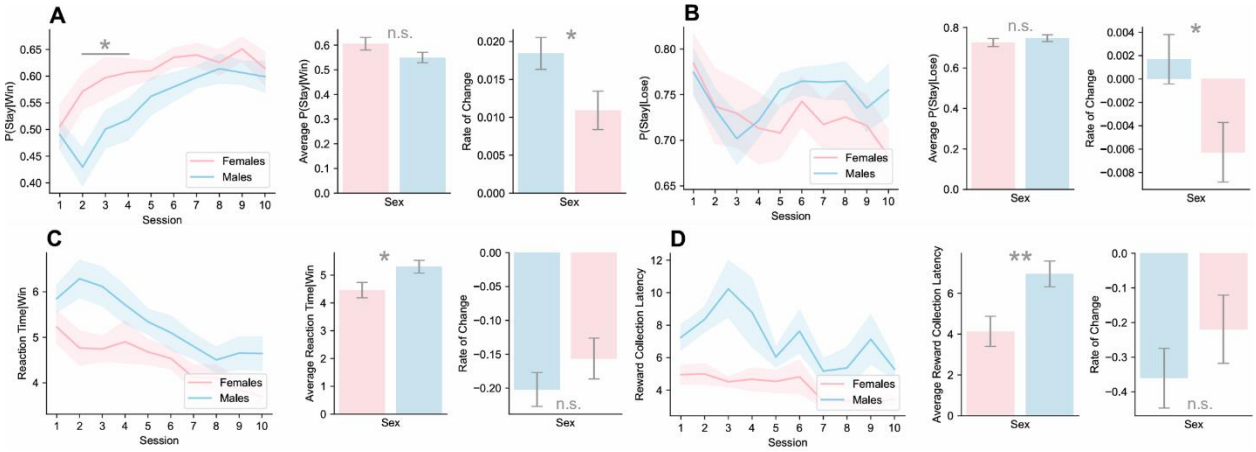


Figure 4. Sex-Differences in Outcome Modulation of Behaviour. A. Win-stay ratios increased at a faster rate in males compared to females (p=0.024) **B.** Lose-stay ratios increased over sessions among males but decreased among females (p=0.017). **C.** Males were slower than females on average to press a lever following a reinforced trial (p=0.012) **D.** Males displayed slower reward collection latency on average (p=0.004).

Conclusion

Wild-type male and female mice were trained to perform a deterministic reversal learning task (DRL) in which one lever is reinforced by the delivery of sweet-condensed milk and illumination of the food port while the other is punished by a cue light. Active and inactive lever locations reverse following 5 consecutive presses on the reinforced lever. While both sexes acquired the task at the same rate, females performed better overall. They earned more rewards per session and were faster at adjusting their behaviour following a reversal by the last session of training. No sex-differences in task engagement where detected. Higher performance in females was in part attributed to better integration of positive reinforcement. Females were faster to react to a cue predicting reward delivery and to make a subsequent decision following reward delivery. While females were able to develop a win-stay lose-shift strategy over time, males persisted in displaying maladaptive lose-stay behaviour. Overall, these results highlight sex-differences in learning to navigate a complex environment in the absence of food-restriction. Males and females displayed distinct development of learning strategies over time and different reactivity to reward delivery.

Chapter 2: Chronic Cocaine Exposure Impairs Decision-Making on a Probabilistic Reversal Learning Task in a Sex and Time-dependent Manner.

Background.

5,8 million cases of Cocaine Use Disorders (CUDs) were reported in 2016 (Degenhardt et al., 2018). Despite the development of behavioural-based treatments, relapse remains a challenge. The relapse rate among cocaine users following treatments is 60-70% (Bisaga et al., 2010). Impaired decision-making is a core feature of CUD. CUD patients display impaired sensitivity to the consequences of their choices, even when working for non-drug associated rewards (Balconi et al., 2014; Barry & Petry, 2008; Kjome et al., 2010; Nigro et al., 2021; Verdejo-Garcia et al., 2007), with worse decision-making impairments predictive of higher relapse rate (Nejtek et al., 2013; Schmitz et al., 2009; Verdejo-Garcia et al., 2014). Better characterization of the altered decision-making processes associated with chronic cocaine exposure is needed to develop more targeted behavioural treatments for CUD.

Computational psychiatry is a novel field aimed at using computational modeling to provide a mechanistic description of behaviour. Reinforcement Learning (RL) models have been instrumental in providing additional insight into the impaired processes in CUD. Clinical work has highlighted a decrease in both negative and positive learning rates in CUD patients compared to controls (Lim et al., 2021), with abstinence selectively remediating impaired positive learning rates (Wang et al., 2020). Rodent models of CUD have demonstrated that lower positive learning rates at baseline predict future increases in cocaine self-administration (Groman et al., 2020a), whereas lower negative learning rates were found to decrease following chronic cocaine exposure (Groman et al., 2020b).

While the aforementioned work demonstrated a link between chronic cocaine exposure and altered processing of negative and positive outcomes, several outstanding questions remain. First, sex has not been investigated as a meaningful variable in mediating the impact of cocaine on reinforcement learning. Sex and gender differences have, however, been reported in the presentation of SUD (Becker, 2012) and baseline decision-making (Kutlu et al, 2020), highlighting the potential for sex-differences in cocaine-affected decision-making. Secondly, the evolution of deficits over the course of cocaine exposure has not yet been investigated. As individuals are continuously navigating the world throughout their cocaine consumption, it is crucial to understand precisely how daily exposure to substances interacts with everyday decisions.

Aims and Hypotheses.

Decision-making of male and female C57BL/6J wild-type mice was assessed on a probabilistic reversal learning task throughout fourteen daily injections of cocaine or saline and following two weeks of forced abstinence. The overarching aim of this experiment is to characterize changes in learning from positive and negative reinforcement across sexes and cocaine administration. Based on previous literature, we expected learning from negative outcomes to decrease over time in a sex-dependent manner.

Results

1. Mice Learn to Perform a Probabilistic Reversal Learning Task.

C57BL/6J wild-type mice (N=28, F=11, 6-7 weeks old) were trained to perform a Probabilistic Reversal Learning (PRL) task. Initially, all mice acquired a lever pressing response on a Fixed Ratio 1 (FR1) schedule reinforced by 10% sweet-condensed milk. Task engagement increased over sessions (**Figure 1.B**; Linear Mixed-Effect Model (LMM), $\beta_{session} = 0.020, p < 0.001$**),** with an average response rate of 54% (SD=12%) by the last session. Following FR1 training, one lever was inactivated for Deterministic Reversal Learning (DRL) task training. The inactive and active lever locations reversed following 5 consecutive presses on the active lever. Responding increased throughout training (**Figure 1.B**; LMM, $\beta_{session} = 0.005, p < 0.001$; overall mean fraction of answered trials=73±33%) as did the probability of pressing the reinforced lever (**Figure 1.C**; LMM, $\beta_{session} = 0.017, p < 0.001$, mean P(Correct Choice) on the last session=49±10%), indicating that mice learned to adapt their behaviour following reversals. However, females performed better than males on average (LMM, $\beta_{sex} = 0.038, p = 0.030$).

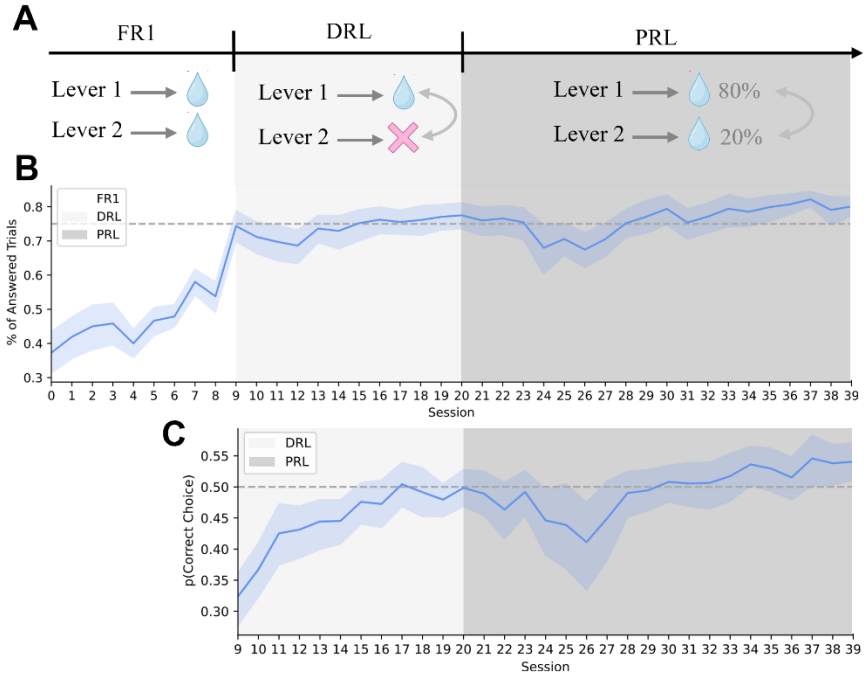


Figure 1. A. Training Protocol. **B.** Task engagement increased over training sessions. **C.** The probability of pressing the reinforced lever increased over training sessions.

Following DRL, decision-making was assessed using a PRL task in which 80% of presses on the "correct lever" were reinforced and 20% of presses on the "incorrect lever" were reinforced. Similar to the DRL task, probabilities reversed following 5 consecutive presses on the correct lever. Mice showed high-task engagement over the last 7 training sessions (**Figure 1C**; mean fraction of answered trials=80±8%, mean number of lever presses=323±83). A subset of mice displayed below chance-level choice accuracy (N=6, F=1; **Figure 2.A and B**). Learners (i.e., mice that pressed the correct lever above chance-levels) were kept for further analyses (N=22, F=10; average P(correct choice) =0.57±0.04). Learners adaptively adjusted behaviour based on the outcome of previous trials and had a higher probability of repeating their choice following rewarded trials, with no effect of sex (**Figure 2C;** two-way ANOVA, main effect of previous outcome: $F(1,52)=53.748$, $p<0.001$; Sex: $F(1, 52)=1.731$, $p=0.194$; Sex x Outcome: $F(1,52)$: 3.244, $p=0.077$). Conversely, non-learners displayed perseverative behaviour regardless of trial outcome (**Figure 2D**; $t(5)=-0.35$, $p=0.741$; not enough females to test for sex-differences). While further

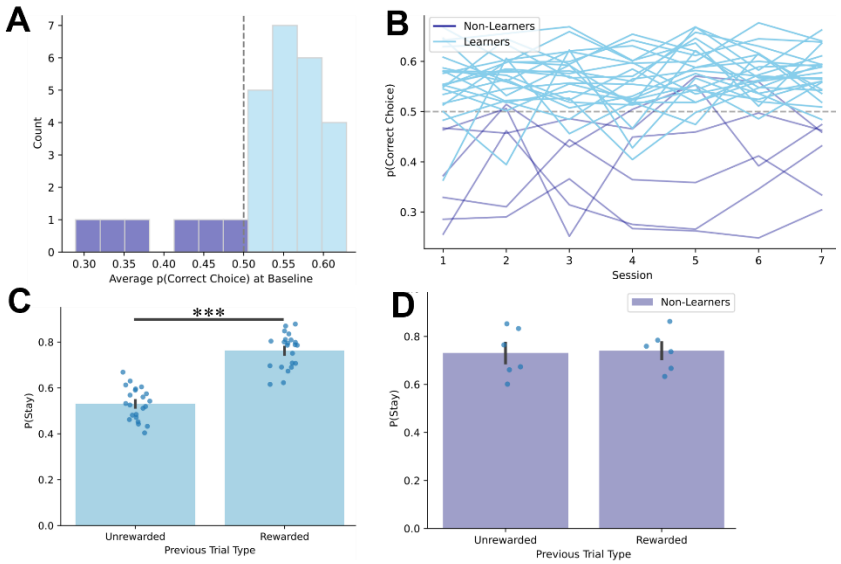


Figure 2. Baseline performances on the PRL task. **A.** Average probability of making a correct choice on the PRL task. A subset of mice (N=6) performed below chance-level. **B.** Probability of making a correct choice during baseline sessions of the PRL task for learners and non-learners. **C.** Learners adaptively adjusted choice behavior based on previous trial outcomes. **D.** Non-learners displayed perseverative behaviour irrespective of previous trial outcomes.

characterization of the behaviour of non-learners could help understand individual differences in learning; the aim of the present experiment is to identify changes in decision-making following chronic cocaine exposure. Non-learners were hence excluded from further analyses to focus on the population of mice that learnt to perform the task and in which changes in decision-making could be observed.

2. Chronic cocaine exposure impairs decision-making

To characterize changes in decision-making throughout chronic cocaine exposure, mice were administered either 20 mg/kg i.p. injections of cocaine (N=10,F=6) or saline (N=10, F=5) for 14 daily sessions **(Figure 3.A).** Injections were administered one hour after testing on the PRL, such

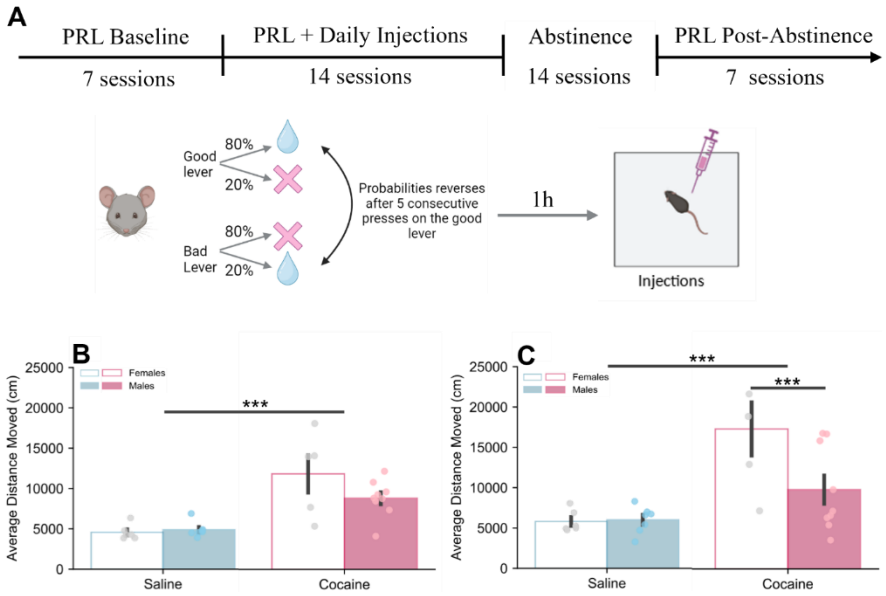


Figure 3. A. Experimental Protocol **B.** The distance moved immediately following cocaine injections increased during the first week of injections. **C.** Females showed greater cocaine-induced locomotion compared to males over the second week of injection.

that decision-making was assessed 23 hours after each injection to avoid the confounding influence of the acute effects of cocaine. Following two weeks of injections, all mice underwent a two-week period of forced abstinence during which both injections and PRL testing stopped. Decision-making was then assessed for one additional week following abstinence to characterize the long-term impacts of chronic cocaine exposure. Cocaine injections acutely increased locomotor activity over 25 minutes following injections in a sex and time dependent manner. Both males and females showed increased locomotion following cocaine injections, with an enhanced effect of cocaine in females compared to males during the second week of injection only (**Figure 3.B and C,** LMM, Drug x Sex x Week: $F(1, 329.05)=18.717$, $p<0.001$; post hoc tests, week 1, cocaine vs saline, $t(24.7)=-5517$, $p<0.001$; week 2, cocaine vs saline, $t(24.6)=-7546$, $p<0.001$; week 2, cocaine females vs cocaine males, $t(24.6)=-7458$, $p<0.001$).

We next assessed the impact of cocaine injections on the navigation of the PRL task. Cocaine injections had no impact on the number of reversals achieved and the total number of rewards earned (**see Supplemental Figure 5**). However, these measures are confounded by fluctuation in task engagement. While the probability of pressing the lever did not differ between drug group, it did fluctuate over time (**Figure 4.C**; Drug x Sex x Time interaction: $F(3, 409.87)=3.142$, $p=0.025$; all between drug group post-hoc tests $p>0.05$). To assess decision-making, we hence computed choice accuracy as the average probability of making a correct choice at each trial. Cocaine injections impaired decision-making in both sexes across all time points (**Figure 4.A;** LMM, Drug

x Time: F(3, 447)=0.432, p<0.001; week 1: t(23.1)=2.289, p=0.031; week 2: t(25.2)=2.682, p=0.012; post-abstinence: t(27.6)=2.764, p<0.001). Males were more accurate than females during

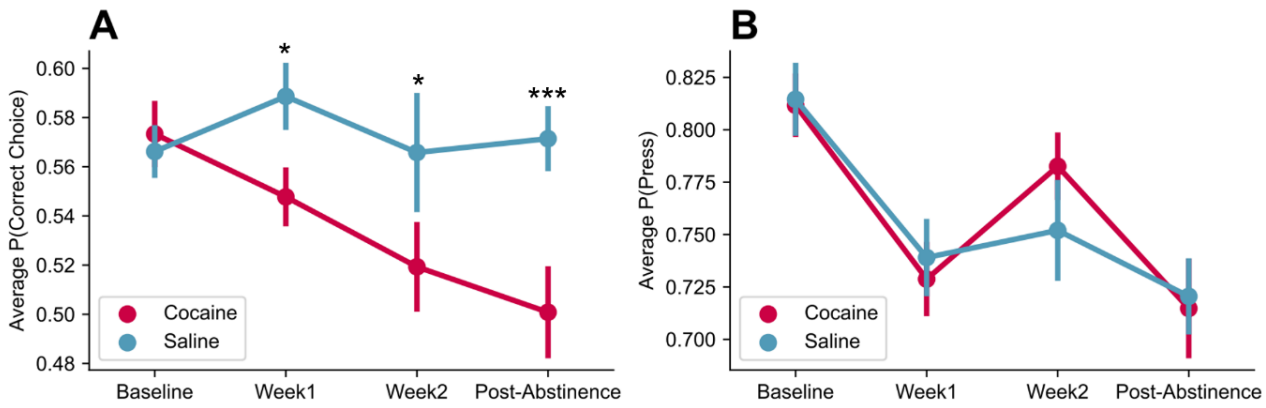


Figure 4. A. Cocaine decreased choice accuracy on the PRL across all time points. **B.** Task engagement fluctuated over time with no difference between drug conditions.

the second week of injections, irrespective of what was injected, highlighting sex-differences (Sex x Time: F(3, 447)=4.819,p=0.003, week 2: t(25.2)=2.43, p=0.042). The decrease in accuracy did not correlate with the locomotor effects of cocaine **(see Supplemental Figure 5**).

3. Chronic Cocaine Selectively Impairs Positive and Negative Reinforcement Learning

We next investigated whether decision-making deficits arose from specific impairments in adjusting behaviour following positive or negative reinforcement, as reflected in the Win-Stay and Lose-Stay ratios, respectively. Cocaine decreased the Win-Stay ratio among males only during the second week of injections (**Figure 5A and B,** LMM, Sex x Time x Drug: F(3, 44.29)=2.9699, p=0.032, t(20)=0.101, p=0.045) with a trend towards significance at abstinence (t(21.1)=2.005, p=0.058). Conversely, the Lose-Stay ratio increased following during the first week of cocaine injection with no significant effect of sex (**Figure 5C,** LMM, Time x Drug: F(3, 465.01)=4.5691,

p=0.004, t(20.6)=-0.067, p=0.029) Cocaine injections hence impacted choice behavior in a time and sex-dependent manner.

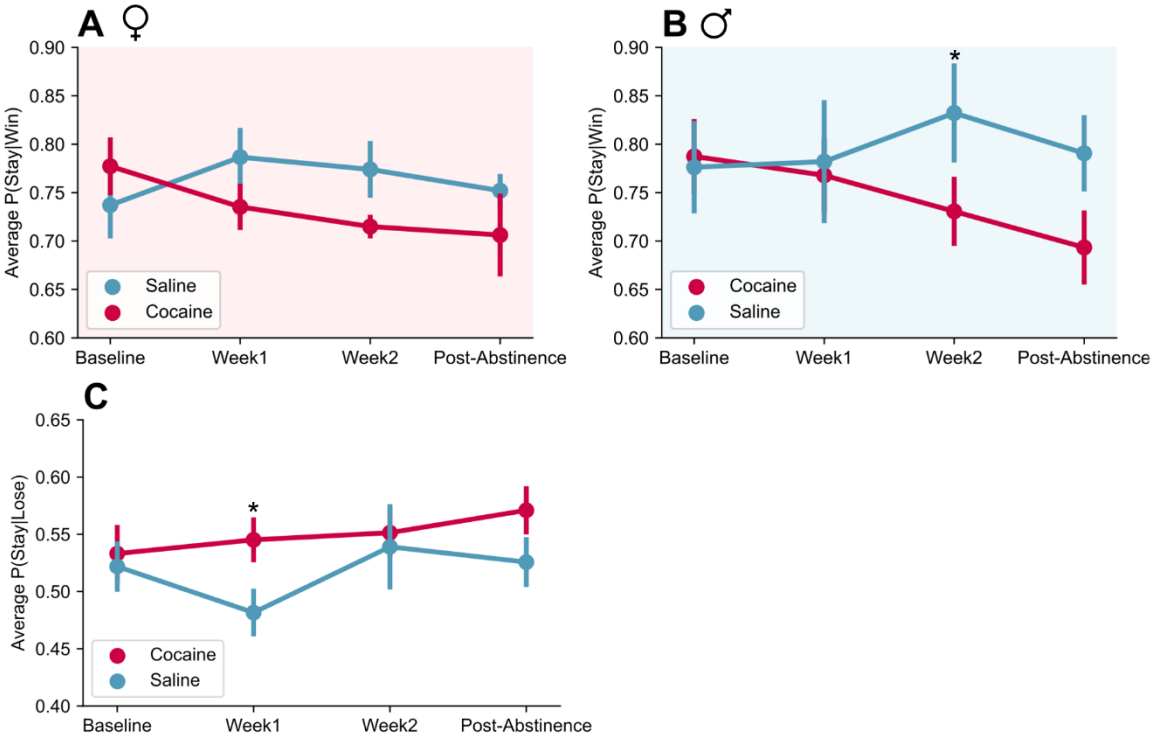


Figure 5. A. Cocaine had no effect on the probability of repeating actions following positive reinforcement in females. **B.** Cocaine-injected male mice were more likely to switch choice following positive reinforcement than saline mice over the second week of injection. **C.** Cocaine transiently increased the likelihood of repeating action despite lack of reinforcement during the first week of injection with no effect of sex.

To characterize whether deficits in adjusting choice based on previous trial outcome could be attributed to impaired integration of recent outcome history, rather than the previous trial only, choice data was fit using a double-Q reinforcement learning model. Sessions from each of the time points were concatenated and modeled independently. Outcome integration was modeled using two learning rates, one for positive outcomes (α+) and one for negative outcomes (α-) (see

Methods). No significant difference between drugs, sex, or time point were detected for the negative learning rate (**Figure 6.B**; LMM, Drug x Sex x Time: F(3,48) = 0.824, p=0.487). However, the positive learning rate decreased in cocaine mice compared to saline at abstinence, irrespective of sex (**Figure 6.A**; LMM, Drug x Time: F(3,48)=3.271, p=0.029; t(41.7) = 2.302,

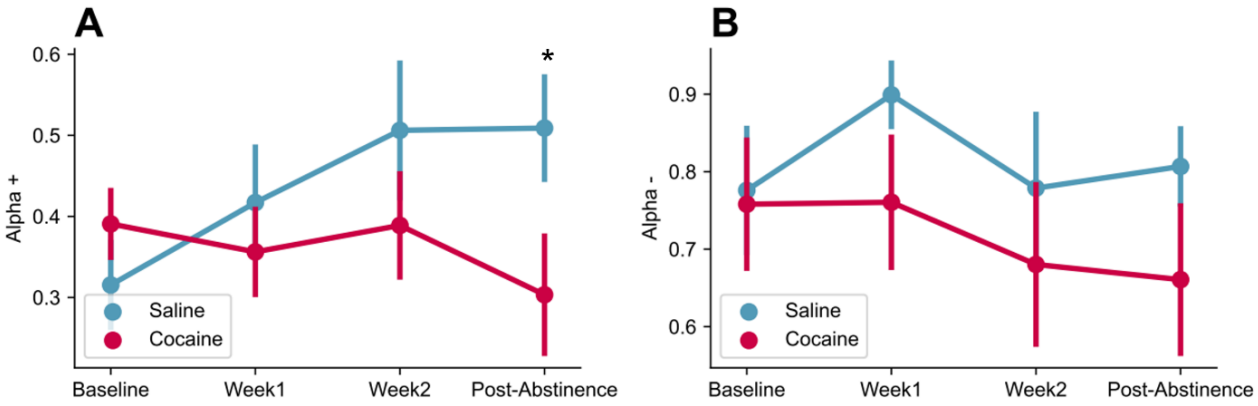


Figure 6. A. Cocaine decreased learning rate for positive outcome following abstinence with no impact of sex **B.** No significant drug nor sex differences were detected in learning rate for negative outcomes.

p=0.039). No significant group differences were detected in the forgetting rate or policy parameters (**see Supplemental Figure 6**).

4. Lower Lose-Stay Ratio Predicts Greater Cocaine-Induced Impairments in Males.

We next asked whether certain decision-making phenotypes made mice more susceptible to cocaine-induced impairments. Baseline task variables on the PRL task for cocaine-injected male and female mice were correlated with the magnitude of change in choice accuracy at each injection time point. Higher probability of making a correct choice at baseline predicted greater impairments during the first week of injections for females (**Figure 7.A**, Pearson coefficient=-0.915, p=0.029) but not males (Pearson coefficient=0.095, p=0.879). Additionally, a lower Lose-Stay ratio at

baseline predicted worse impairments following abstinence in males (**Figure 7.B,** Pearson coefficient=0.896, p=0.039) but not females (Pearson coefficient=0.167, p=0.723).

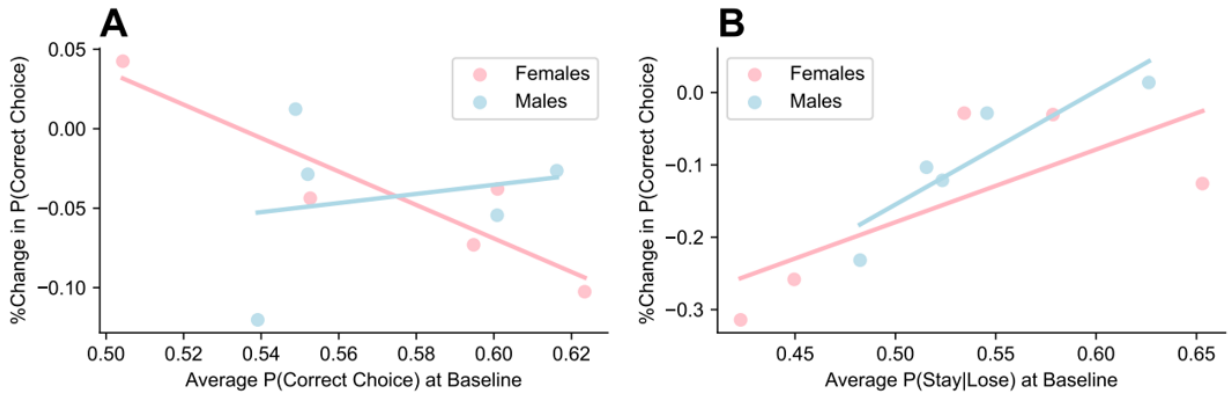


Figure 7. A. Higher choice accuracy at baseline predicts greater impairments during the first week of cocaine injections among females but not males. **B.** Lower lose-stay ratio at baseline predicts greater impairments follow abstinence from cocaine injections among males but not females.

Conclusion

C57BL/6J wild-type mice (N=28, F=11, 6-7 weeks old) were injected with cocaine or saline daily for fourteen sessions. Decision-making was assessed before, during, and after cocaine exposure using a PRL task. Cocaine impaired decision-making in both male and female mice during the first week of injections, with effects persisting across two weeks of abstinence. Male, but not female, mice showed a specific decrease in P(Stay|Win) following the second week of cocaine injections, while both sexes exhibited a transient increase in P(Stay|Lose) during the first week. Reinforcement learning (RL) modeling revealed decreased learning rates from positive outcomes in cocaine-treated vs. saline-treated mice after abstinence, regardless of sex. This research demonstrates the direct impact of chronic cocaine exposure on flexible decision-making, leading to a lasting reduction in choice accuracy following injections. Decreased sensitivity to

positive outcomes was observed in males during the second week of injections and in both sexes after abstinence. Conversely, cocaine transiently increased perseveration.

1. Sex-Differences in Decision-Making.

C57BL6J wild-type mice were trained on a deterministic reversal learning (DRL) task for 10 sessions to compare learning trajectories and decision-making strategies in ad libitum fed male and female mice. While both sexes acquired the task at the same rate, females displayed better task performance than males and adapted faster to changes in contingencies. Females were more likely to persevere with rewarded choices, had faster reward collection latencies, and were faster to press the lever again following reward delivery, indicating higher sensitivity to positive reinforcement. No sex-differences were detected in choice behavior or reaction times following non-reinforced trials. Males and females showed different learning trajectories over time. Across sessions, females decreased their likelihood of persevering with non-reinforced levers, whereas males did not. Conversely, males displayed faster increases in the likelihood of choosing the same lever following reinforced trials, despite an overall lower rate than females. These results highlight sex-differences in decision-making strategies, with females showing higher sensitivity to positive reinforcement and better adaptive responses to negative outcomes over time.

1.1 Females Perform Better than Males on a Reversal Learning Task.

In the DRL task, females earned more rewards than males and showed better adjustment to reversals, even in the first session, which is indicative of better behavioural flexibility. In contrast, a previous study found that female mice displayed similar performance to males when reward probabilities were deterministic in a two-armed bandit task, with increases in performance only observed when rewards deliveries were rendered probabilistic (Chen, Ebitz, et al., 2021). A crucial difference in the design of that task was the presence of cues to indicate the location of the high probability option, whereas the levers in our task were identical. This difference suggests that sex differences in decision-making may only arise when there is uncertainty in choice outcomes, whether from probabilistic reward delivery or uncued choices and reversals. Relatedly, females

are better at learning to avoid shock when it is delivered in a probabilistic but not deterministic manner (Chowdhury et al., 2019), and they prefer smaller, safer rewards over larger ones that come with the risk of footshock, highlighting sex-differences in resolving uncertainty and conflict (Orsini et al., 2022).

Why are females better at making decisions in uncertain environments? Our results highlight a differential integration of outcomes. Males' reaction times were not influenced by previous trial outcomes, and males showed a lower initial rate of win-stay behavior than females. Additionally, males did not decrease their lose-stay behavior over training sessions. Conversely, females adaptively adjusted choice behaviour depending on previous trial outcomes, showed faster reward collection latencies, and were faster to reengage in the task after reward delivery. These behavioral results mirror what has been observed in reinforcement learning models fit to the behavior of mice on a two-armed restless bandit task, which highlight lower learning rates in males (Chen, Knep, et al., 2021). Furthermore, the higher win-stay behavior in females suggests increased sensitivity to positive reinforcement. Likewise, women show better accuracy when learning from positive but not negative feedback (Evans & Hampson, 2015). Faster reward collection latency among female mice arose during FR1 training. Since a light cue indicated the delivery of reward, this sex-difference could reflect differences in Pavlovian rather than instrumental learning. Supporting this idea, when head-fixed mice were trained on a Pavlovian task in which audiovisual cues predicted reward delivery, females were faster than males to retrieve rewards (Lefner et al., 2022).

1.2 Gonadal-Hormones and Decision-Making.

Previous work has found sex differences in decision making, such as increased risk-avoidance in females (Islas-Preciado et al., 2020; Liley et al., 2019; Orsini et al., 2016). We found that female

mice were better than males in adjusting to contingency reversals and learning from previous trial outcomes. Levels of gonadal hormones can act as a potential mediator of such sex-differences in decision-making. Estrogens and progesterone are the main female gonadal hormones. Their respective levels fluctuate throughout the estrus cycle, the rodent equivalent of the menstrual cycle, and the highest levels are seen during the proestrus and estrus phases (Smith et al., 1975). Conversely, testosterone levels in males primarily fluctuate within a day, instead of the 96 hour-long estrus cycle in females; levels are highest in the morning and decline throughout the day (Heywood, 1980). Importantly, all of these hormones – progesterone, estrogen and testosterone – can be synthesized locally in the brain and are present in both sexes, albeit at different concentrations (Orsini et al., 2022).

Can estradiol play a role in faster integration of outcomes and increased reversal learning? While natural fluctuations of the estrus cycle have not been associated with changes in decision-making (Georgiou et al., 2018; Orsini et al., 2016; Van Den Bos et al., 2012), estradiol treatment following ovariectomy (OVX) in female rats and mice decreased risk-taking (Orsini et al., 2021) and improved working and spatial memory (J. Daniel, 2004; J. M. Daniel et al., 2006), reversal learning (Arad & Weiner, 2012), and renewal of a conditioned response following extinction (Anderson & Petrovich, 2015). In males, estradiol treatment following orchiectomy (ORX) similarly decreases risk-taking (Orsini et al., 2021). While our data does not address the potential impact of hormone fluctuations on female learning, since we did not track the estrus cycle, higher performance among females relative to males may be linked to overall higher estradiol levels.

Conversely, testosterone has been reported to increase rats' aversion to uncertainty. Testosterone treatment in intact males decreased preference for large, low probability rewards and biased choices toward smaller, more certain rewards (Wallin et al., 2015; Wallin-Miller et al.,

2018). However, when a guaranteed large reward was associated with a probability of foot shock, testosterone treatment increased choice for the riskier option (Cooper et al., 2014). ORX decreased risky choice on a similar task (Orsini et al., 2021). It is possible that testosterone may decrease sensitivity to physical punishment while increasing aversion to uncertainty. We found decreased adaptation to reversal in males, possibly because the higher testosterone levels in males promoted choice for a 'safer' option that previously led to reward delivery and an aversion to adapting behaviour towards a less certain option.

1.3 Sex-Differences in the Neurobiological Substrate of Decision-Making.

Decision-making requires the interaction of multiple brain regions across the mesocorticolimbic system, including the prefrontal cortex (PFC) and the nucleus accumbens (NAC), which are both densely innervated by dopaminergic (DA) projections from the ventral tegmental area (VTA**)** (Orsini et al., 2022). VTA DA activity putatively encodes a reward prediction error (RPE) to drive reinforcement learning (Schultz et al., 1993). VTA DA hyperactivity impairs reversal learning and increases risk-taking in a probabilistic discounting task (Verharen et al., 2018). Gonadal hormones have been reported to regulate VTA DA neuron activity, with female mice showing increased basal activity of DA neurons in the ventral tegmental area (VTA) during the estrous phase (Calipari et al., 2017). DA regulation may hence be key to understanding sex-differences in decision-making. The following section provides a brief overview of the involvement of regions downstream of the VTA in decision-making, specifically the NAC and PFC, in light of reports of sex-differences in these regions.

1.3.1. Nucleus Accumbens.

We found faster integration of outcomes and higher sensitivity to positive reinforcement among females. Value-updating (or outcome integration) has been hypothesized to be mediated by post-

synaptic targets of DA in the NAC. The NAC has conventionally been divided into two main dopamine targets: D1- and D2-receptor-expressing medium spiny neurons (D1 and D2 MSNs). D1 and D2 MSNs display distinct physiological and cellular properties and have different projection targets (Gerfen & Surmeier, 2011; Tepper et al., 2010). The D1 receptor has a low affinity to dopamine and thus may only be activated by phasic DA release (Marcellino et al., 2012). In contrast, D2 receptors have a high affinity to dopamine, and tonic DA levels are sufficient to activate them (Bariselli et al., 2019) This difference has inspired theories in which D1 and D2 MSNs are assigned dedicated and distinct functions, namely modulating positive and negative learning rates, respectively (Frank et al., 2004; Romero Pinto & Uchida, 2023).

Both systemic injections and NAC microinfusions of D2 receptor agonists impair reversal learning (Alsiö et al., 2019; Boulougouris et al., 2009; Haluk & Floresco, 2009) and decrease risk-taking when a large reward is associated with the risk of footshock (Mitchell et al., 2014; Simon et al., 2011). Inhibition of D2 MSNs in the NAC similarly impairs reversal learning (Macpherson et al., 2022). Optogenetic stimulation of these neurons during deliberation periods decreases risk-taking (Zalocusky et al., 2016), whereas their inhibition increases it (Truckenbrod et al., 2023). Hence, modulation of D2 receptor activity and D2 MSN activity both impair reversal learning and risk-taking. D2 receptor and MSN activity have been hypothesized to regulate learning from negative outcomes. Individuals with lower D2 receptor density display reduced learning from negative outcomes (Cox et al., 2015), and D2 receptor agonists decreased negative but not positive learning rates in mice (Alsiö et al., 2019). Furthermore, calcium imagining of D2 MSNs in the NAC revealed that these neurons encoded for previous losses on a decision-making task and biased future choices (Zalocusky et al., 2016). Sex-differences in decision-making could hence arise from neurobiological differences in NAC D2 MSN activity. These neurons express estradiol receptors,

and estradiol treatments augment D2 receptor binding (Orsini et al., 2022; Yoest et al., 2018). Notably, females display increased D2 receptor density in the striatum compared to males (Hasbi et al., 2020; Orendain-Jaime et al., 2016).Testosterone levels also regulate D2 MSN activity, but testosterone treatments decrease D2 receptor density in the NAC (Orsini et al., 2022). While testosterone treatments decrease choice for larger but probabilistic rewards, a D2 receptor agonist reversed the effect, highlighting a potential mechanism by which testosterone impacts decision-making through D2 MSN activity (Wallin-Miller et al., 2018).

Conversely, D1 receptor density in the striatum correlates with learning from positive outcomes in healthy individuals (Cox et al., 2015), and polymorphisms in the DARPP32 gene, which have been linked to D1 MSN function, predict learning from positive outcomes (Frank et al., 2007). While infusion of either a D1 receptor antagonist or agonist in the NAC does not impact reversal learning, a D1 antagonist impaired set-shifting (i.e., the ability to switch decision-making strategy (Haluk & Floresco, 2009). Sex-differences have been reported in D1 receptor density in the NAC, with higher D1 receptor expression in male rats compared to females (Andersen et al., 1997; Hasbi et al., 2020).

1.3.2. Prefrontal Cortex.

Patients with PFC lesions display deficits on the IGT (Bechara et al., 1994; Fellows, 2004). Distinct PFC subregions have been reported to mediate diverse aspects of decision-making. The rodent PFC can be decomposed into medial orbitofrontal cortex (mOFC), lateral orbitofrontal cortex (lOFC), infralimbic cortex (IL), and prelimbic cortex (PL). Contrasting results have been reported on the involvement of two OFC subdivisions in decision-making. Inactivation of the medial OFC (mOFC) and lateral OFC (lOFC) in male rats increased and decreased risk-taking, respectively, across studies in which the risky-option was defined as a delivery of larger but

probabilistic reward (Stopper et al., 2014), a delayed reward (Mar et al., 2011), or the risk of being physically punished (Orsini et al., 2015). Similarly, mOFC lesions improved adjustment to contingency reversals, while lOFC lesions impaired it (Mar et al., 2011). With respect to the IL and PL, combined inactivation of these regions caused deficits on a rodent gambling task (Zeeb et al., 2015). The PL has been specifically implicated in behavioural flexibility in dynamic environments. PL inactivation only increased risk-taking when the probability of risk increased over the session (i.e., an increasing probability of outcome omission or footshock over trials) (Orsini et al., 2018; St. Onge & Floresco, 2010). Overall, research points to distinct roles of PFC subregions in mediating decision-making. In addition, cells expressing different types of receptors contribute differently to decision-making. Infusion of a D1 receptor antagonist into either the PL or the mOFC decreases probabilistic risk-taking, whereas a D2 receptor antagonist increases it (Jenni et al., 2021; St. Onge et al., 2011). Similarly, a D1 antagonist impaired performance on a probabilistic reversal learning, while a D2 antagonist improved it (Jenni et al., 2021), pointing to a role in behavioural flexibility. Further cell-type classifications could explain the different contributions of the distinct anatomical subregions of the PFC. For instance, cell types with similar transcriptomic profiles in the mouse PFC permitted higher accuracies in decoding future decisions than did cells classified according to projection target (Lui et al., 2021). The aforementioned work highlights the involvement of PFC regions in weighing outcomes to guide decisions. A distinct conceptual approach to studying the role of the PFC in decision-making is based on the theory that PFC encodes task state. Reinforcement learning (RL) theory posits that agents take decisions by computing the values of an abstract state. In simple task structures, a state can be a visual cue indicating the location of a reinforced lever. However, a state can be more complex, such as an abstract location in space. Efficient representation of the task structure is crucial to learning a new

task (Wilson & Niv, 2012). Works in humans and monkeys have reported encoding of task state in neuronal activity in the OFC (Schuck et al., 2016; Wilson et al., 2014) and dLPFC (Bartolo & Averbeck, 2020).

Critically, all the aforementioned works was conducted in male subjects, despite evidence that gonadal hormones regulate DA neurotransmission in the PFC. Female rats show higher basal mPFC DA levels with higher estradiol levels (Dazzi et al., 2007). OVX decreases mPFC DA levels (Kokras et al., 2018), and estradiol injections following OVX both increases DA levels and upregulates DA receptors (Jacome et al., 2010; Sárvári et al., 2014). Testosterone similarly modulates PFC activity and DA transmission. Androgen receptors (AR) are expressed on VTA dopamine neurons that project to the mPFC as well as on mPFC glutamate neurons that project to the VTA (Aubele & Kritzer, 2012; Kritzer & Creutz, 2008), and there is greater AR expression in VTA-PFC neurons in males compared to females (Kritzer & Creutz, 2008). In humans, functional magnetic resonance imaging (fMRI) revealed increased activity in the OFC during anticipation of reward in a decision-making task that was modulated bythe phase of the menstrual cycle (Bayer et al., 2013; Dreher et al., 2007)

1.4. Considerations for Including Sex in Decision-Making Research.

Our results demonstrate clear sex-differences in the acquisition of reversal learning as well as behavioural differences following positive and negative outcomes. Other mechanisms beyond gonadal hormones may mediate sex differences in value-based decision-making For instance, genetically modified mice made to express ovaries or testes with both XX and XY genotypes have uncovered differences between XX and XY genotypes, regardless of gonadal status, in motivation (Seu et al., 2014), habit formation (Barker et al., 2010; Quinn et al., 2007) and performance on the

PRL task (Aarde et al., 2021). Sex-differences may also arise from differential environmental interactions (McCarthy & Arnold, 2011).

Despite the 2016 NIH mandate for the consideration of sex as a biological variable and reports of sex-differences in multiple disorders such as substance use disorders (Becker et al., 2017), research still shows a clear bias toward male samples (Nunamaker & Turner, 2023). Future research will have to include male and female samples and properly analyze and test for the presence of sex-differences.

2. Chronic Cocaine Exposure Induces Reinforcement Learning Deficits.

With daily injections of 20 mg/kg cocaine, we found that male and female wild-type mice displayed impairments in probabilistic reversal learning, with deficits persisting across two weeks of forced abstinence. Decision-making deficits were accompanied by a transient increase in lose-stay probability during the first week of injection. Male but not female mice showed a decrease in win-stay behavior during the second week of injections. RL modeling further revealed decreased positive learning rates following abstinence in both sexes.

2.1. Cocaine Impacts the Sensitivity to Positive and Negative Outcomes in a Time Dependent Manner.

Previous work has reported increased lose-stay behaviour in male rats following cocaine self-administration (Groman et al., 2020a; Zhukovsky et al., 2019). We not only extend these findings to females, but further show that deficits are transient and can be ameliorated through continuous training on the PRL task. Increased lose-stay behavior reflects perseverative behaviour despite lack of reinforcement and has been associated with a decreased sensitivity to negative outcomes. In contrast to impaired lose-stay behaviour, we do not report a decrease in the negative learning rate parameter. However, RL modeling assumes that animals use the history of reinforcement to

progressively update action-values. Conversely, the lose-stay measure only utilizes the outcome of the previous trial to guide the next decision. It is possible that we did not detect changes in negative learning rate because animals relied on the immediate previous negative outcome to guide their decisions. In fact, a subset of mice had negative learning rate values close to 1, indicative of immediate updating of values. The discrepancies between learning rates and lose-stay behavior may additionally be due to their non-linear relationship. Simulations of behavior using RL modeling have shown a higher lose-stay ratio at both low and high learning rate values, with the lowest ratios observed at intermediate learning rates (Iyer et al., 2020). Therefore, the increase in lose-stay observed during the first week of cocaine injections could correspond to both high and low learning rates, making it difficult to detect significant differences in learning rates. However, in our experiment, we did not observe a non-linear relationship between the lose-stay ratio and negative learning rates. During the first week of injections, the lose-stay ratio and negative learning rate were negatively correlated (see Supplementary Figure 7), indicating that a higher lose-stay ratio reflected lower negative outcome integration.

While the decreased outcome sensitivity following cocaine injections mirrors findings from humans showing decreased loss avoidance in humans (Lim et al., 2021), the transient nature of it is novel. As the effect disappeared during the second week of injections, we cannot attribute it to an abstinence effect. It is possible that the sensitivity to negative outcomes normalizes with further exposure to cocaine or with further training on the task. Interestingly, one study in humans reported that cocaine users remain worse than controls at avoiding shocks despite extensive training (Ersche et al., 2008). These different results may relate to differences in learning from a lack of reinforcement vs explicit physical punishment.

During the second week of injections, lose-stay behaviour normalized in both sexes, but the win-stay ratio decreased within males only. Decreased win-stay had been reported to both predict future cocaine escalation in rats (Groman et al., 2020a) and to change following cocaine self-administration (Groman et al., 2020b). Our findings support a selective decrease in win-stay arising as a consequence of cocaine injections.

2.2. Behavioural Predictors of RL deficits.

Our experiment highlights that certain behavioral phenotypes at baseline predict a greater decrease in the probability of earning a reward following cocaine injections, and this effect is sex dependent. Among males, a lower lose-stay ratio at baseline predicted a greater decrease in the probability of reward delivery following abstinence. In contrast, previous work has shown that lower integration of positive, but not negative, outcomes predicts future escalation in cocaine self-administration in male rats (Groman et al., 2020b). A possible explanation for these discrepancies is that in our experiment, cocaine administration is controlled, so we measure the susceptibility to developing reinforcement learning (RL) deficits for a given cocaine dose, whereas Groman et al. (2020b) measure the susceptibility to escalating use. Thus, while lower positive outcome integration predicts greater escalation in cocaine use, we show that lower negative outcome integration predicts greater RL deficits with chronic cocaine exposure. On the other hand, in females, a higher probability of reward delivery at baseline predicted a greater decrease after a week of injections. Although the exact measures differ, better navigation of the task for both sexes—either by being more likely to earn rewards or by being less perseverative following losses—predicted greater future deficits. One interpretation is that to become worse at the task, you first need to have mastered it. Mice with higher choice accuracy have more room to decline, whereas for those already performing poorly, the decrease may be negligible.

2.3. Changes in Dopaminergic Signaling Might Underlie RL Deficits.

As reviewed in section 1.3, dopamine is a key neuromodulator in mediating decision-making, and changes in dopaminergic signaling due to substance exposure may underlie the decrease in flexible decision-making reported here with chronic cocaine injections. Substances of abuse modulate DA signaling in the NAC. Acute exposure to drugs of abuse leads to increased extracellular DA levels (Imperato & Di Chiara, 1985). Interestingly, chronic intake of cocaine leads to long-term neural adaptation in DA signaling and dampens dopaminergic tone. Stimulant users show blunted DA release in response to natural rewards (Volkow et al., 2007), with abstinence remediating that effect and increasing neural responses to positive prediction errors (Wang et al., 2019). Cocaine further disrupts the balance in activity between D1 and D2 neurons in the NAC. Stimulant abuse is associated with a specific downregulation of D2 receptors in the striatum (Everitt et al., 2008; Nader & Czoty, 2005; Volkow et al., 2001). At a molecular level, however, D1 neurons in the shell region of the NAC showed decreased intrinsic excitability and increased synaptic input following chronic cocaine injections, with no change in D2 neurons (Kim et al., 2011). This disruption of the balance of D1 and D2 neurons in the striatum following chronic cocaine injection may underlie cocaine-induced decision-making deficits. Downregulation of D2 receptors and resulting increases in D2 neuron activity may contribute to the enhancement in negative learning rate among stimulant users, whereas the reported decrease in D1 neuron excitability may dampen positive learning rates. Manipulation of D2 receptor activity in patients with SUDs supports this theory. A D2 agonist but not antagonist decreased punishment-driven learning among patients with SUDs, alleviating their enhanced sensitivity to negative reinforcement (Kanen et al., 2019).

2.4.Beyond a D1/D2 Dichotomy?

While some support has been found for the D1/D2 theory of positive/negative reinforcement learning, this framework might be an incomplete view of the NAC. Contrasting findings have been found for the role of NAC D1 MSNs in mediating substance use. NAC D1 MSN activity can induce relapse or extinction depending on their projection target (Gibson et al., 2018), and D1 MSN stimulation can promote both reward and aversion depending on the stimulation parameters (Soares-Cunha et al., 2020). Moreover, a molecularly defined subtype of D1 MSN suppresses cocaine self-administration instead of promoting it (Zhao et al., 2022). Similar challenges arise with the consideration of D2 MSNs as decreasing substance-induced behaviour. D2 receptor agonists and antagonists both decrease avoidance behaviour in patient with SUDs, despite eliciting opposite physiological effects (Lim et al., 2021). Furthermore, the D1/D2 dichotomy ignores the spatial heterogeneity of the NAC. The core and shell divisions of the NAC have distinct molecular features (Brimblecombe & Cragg, 2017). MSNs in similar subregions project to different targets (Kupchik et al., 2015) and have different neural functions (Al-Hasani et al., 2015). Redefining cell subtypes to go beyond the D1/D2 dichotomy might help to gain a better understanding of the neurobiological substrate of cocaine induced impairments. Indeed, transcriptomic profiling after acute cocaine administration in rats showed recruitment of distinct MSNs subpopulations (Savell et al., 2020).

3. Sex Modulates the Relationship Between Cocaine and Behaviour

3.1. Sex-Differences in Cocaine-Induced Reinforcement Learning Deficits.

While cocaine transiently decreased lose-stay behavior and, more enduringly, overall choice accuracy similarly in both sexes in our experiment, males but not females showed a long-term decrease in win-stay. We also found higher sensitivity to positive reinforcement learning among

females in drug-naïve decision-making, similar to research in humans (Evans & Hampson, 2015), which raises the possibility that females may be more resistant to reductions in positive reinforcement. Sex-differences in the impact of cocaine administration on the dopaminergic system might underlie this effect. While 5 days of cocaine injections shifted the balance of NAC shell D1 and D2 neuron excitability toward higher D2 excitability in both sexes, this effects was underlied by a decrease in D1 neuron excitability in males and an increase in D2 neuron excitability in females, with greater effects seen in estrus compared to diestrus (Chapp et al., 2024).

3.2. Sex-Differences in the Locomotor Effects of Cocaine.

Acute injections of 20 mg/kg cocaine increased locomotor activity in both male and female mice over two weeks of injections. However, increases in locomotion were enhanced in female mice over the second week of injections compared to males. Increases in locomotor activity following stimulant exposure are enhanced by repeated administration, a behavioural phenomenon referred to as locomotor sensitization (Liu et al., 2018). The development of locomotor sensitization has been used to highlight the presence of long-term neural adaptation with repeated substance exposure, most prominently in the dopaminergic system (Liu et al., 2018). Interestingly, female rats have been reported to display increased locomotor sensitization compared to males (Becker et al., 2016), an effect paralleled with increased spine density in the NAC among females (Wissman et al., 2011). As such, enhancement in the locomotor effects of cocaine may arise from sex dependent neural adaptations.

4. Conclusion.

Chronic cocaine use is associated with poor cognitive flexibility and impairments in reinforcement learning (RL). While sex-differences in learning and decision-making have been reported previously, its impact on cocaine-induced RL deficits remain poorly understood. Here, we first compared the learning strategies of male and female mice throughout the acquisition of a deterministic reversal learning task, and then assessed the impact of chronic cocaine exposure on the navigation of a probabilistic reversal learning task. We report sex-differences in learning. Female mice earned more rewards than males overall and displayed better integration of positive reinforcement. Additionally, while females were able to develop a win-stay lose-shift strategy over time, males persisted in displaying maladaptive lose-stay behaviour. Chronic cocaine exposure impaired decision-making in both sexes by decreasing the probability of making a correct choice on a probabilistic reversal learning task. Positive evidence integration, measured by the win-stay ratio, decreased during the second week of injection in males only. Positive learning rates were decreased following forced abstinence in both sexes. On the other hand, both sexes displayed a transient decrease in the sensitivity to negative outcomes during the first week of injections, reflected in a decrease in the lose-stay ratio but not negative learning rate. Our experiments extend previous research on cocaine effect on decision-making to both sexes. More research on the impact of sex-characteristics on substance use is needed to develop treatments for SUDs adapted to each individual.

Supplementary Materials

Methods: Reinforcement Learning Modeling.

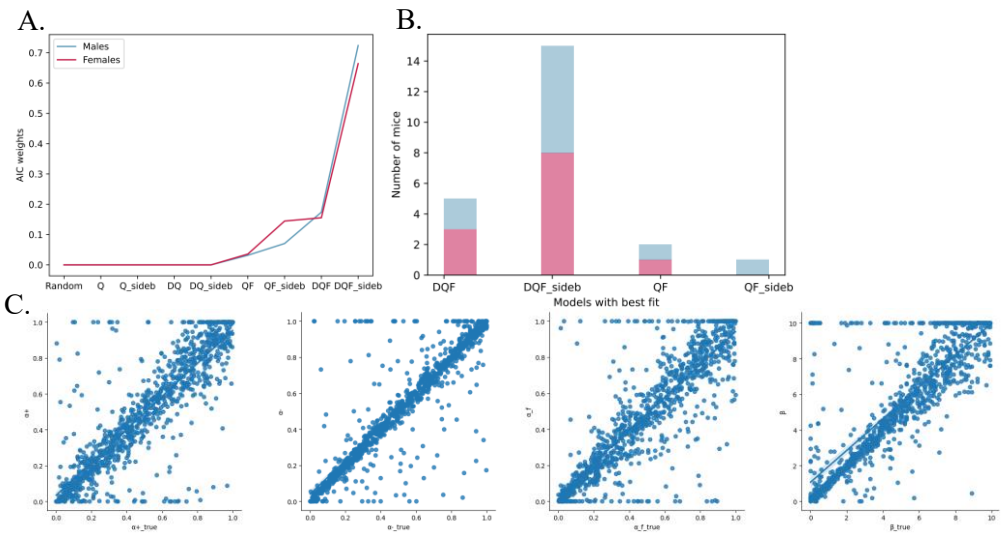
Model Comparisons. 10 variants of reinforcement learning models were fit to choice data of each mouse separately. Supplementary Table 1 describes the parameters used and their interpretation. The models were compared using the Akaike Information Criteria (AIC) weights. Double reinforcement learning with gradual forgetting was found to fit data best on average for both males and females (Sup. Figure 1A). A subset of mice shows better fit using single learning rate parameter (Sup. Figure 1B), in which case the positive and negative learning rates were set to the same value for parameter comparisons. Similarly, if mice did not show additional side bias, the side bias parameter was set to 0.

Model Name	Abbreviation	# Param	Parameters
Random	Random	1	b = tendency to press left
Win-Stay/Loose-Shift	WSLS	1	ε = tendency of win-stay/loose-shift
Q-learning	Q	2	α = learning rate β = exploration rate
Double Q-learning	DQ	3	α^+ = learning rate for rewarded trials α^- = learning rate for non rewarded trials β = exploration rate
Q learning with choice kernel	QCK	4	α = learning rate α_c = learning rate for choice kernel β = exploration rate β_c = exploration rate for choice kernel
Double Q-learning with choice kernel	DQCK	5	α^+ = learning rate for rewarded trials α^- = learning rate for non rewarded trial α_c = learning rate for choice kernel β = exploration rate β_c = exploration rate for choice kernel
Q-learning with choice kernel balanced	QCKe	5	α = learning rate α_c = learning rate for choice kernel β = exploration rate β_c = exploration rate for choice kernel η = tendency to rely on reward history rather than action history
Double Q-learning with choice kernel balanced	DQCKe	6	α^+ = learning rate for rewarded trials α^- = learning rate for non rewarded trials α_c = learning rate for choice kernel β = exploration rate β_c = exploration rate for choice kernel η = tendency to rely on reward history rather than action history
Q-learning with forgetting	QF	3	α = learning rate α_f = forgetting rate β = exploration rate
Double Q-learning with forgetting	DQF	4	α^+ = learning rate for rewarded trials α^- = learning rate for non rewarded trial β = exploration rate

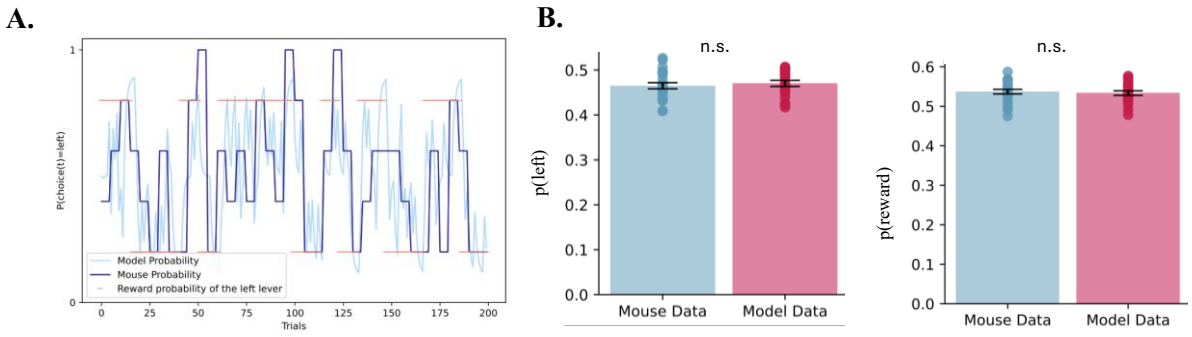
Supplementary Table 1. Parameter descriptions of all RL models fit on choice data.

Parameter Recovery. Parameters estimations were performed via minimization of the negative log likelihood of the data. To validate the parameter recovery method, we generated 1000 choice data from known, randomly generated parameters and plotted recovered over true parameters (Sup. Figure 1C).

Model Validation. To assess the model with best fit accurately described choice data, we used the recovered parameter of each mouse to simulate performance on the PRL task. We then compared choice output from the RL generated data and from mouse data (Sup. Figure 2B).

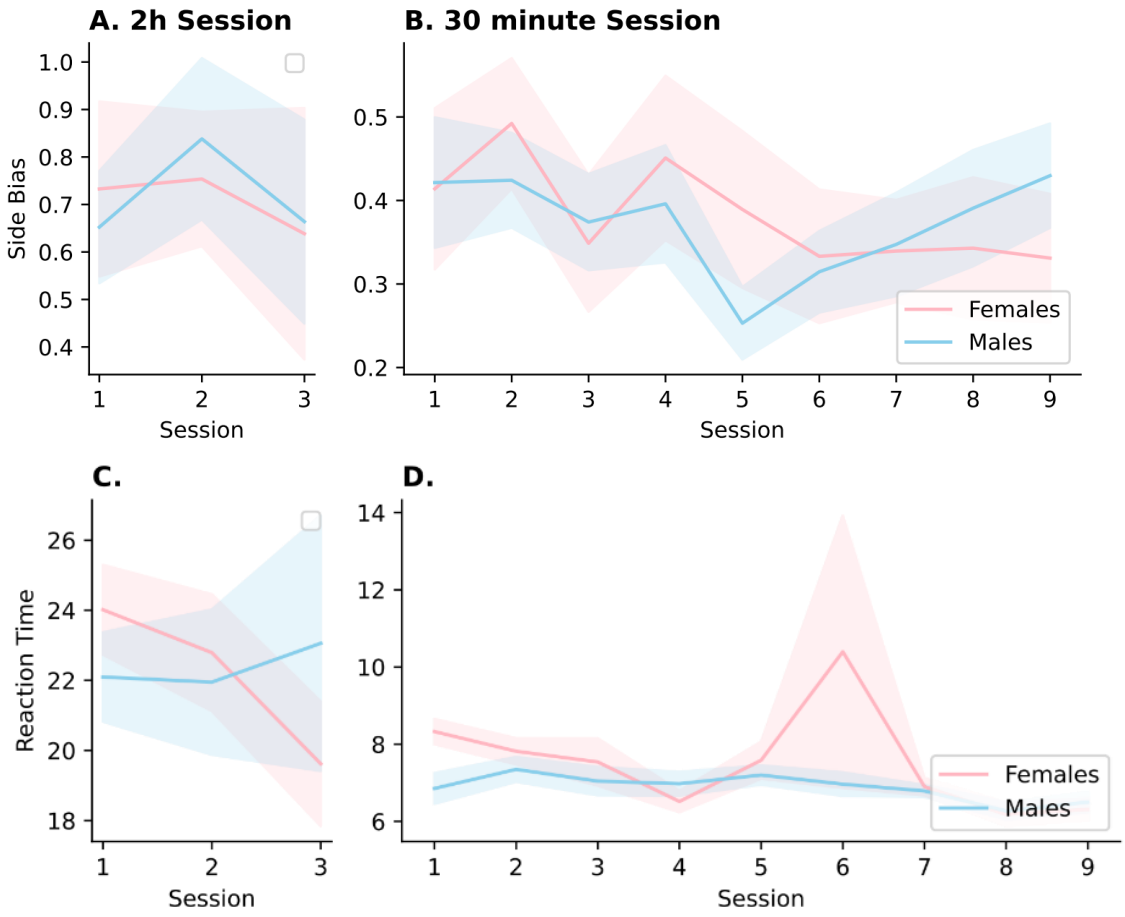


Supplementary Figure 1. **Model Comparison and Parameter Recovery. A.** Average AIC weight values across males and females for all 10 models compared. **B.** Distribution of model with best fit per mouse for males (blue) and females (pink). **C.** Parameter recovery performed on double reinforcement learning with forgetting over 1000 iterations.



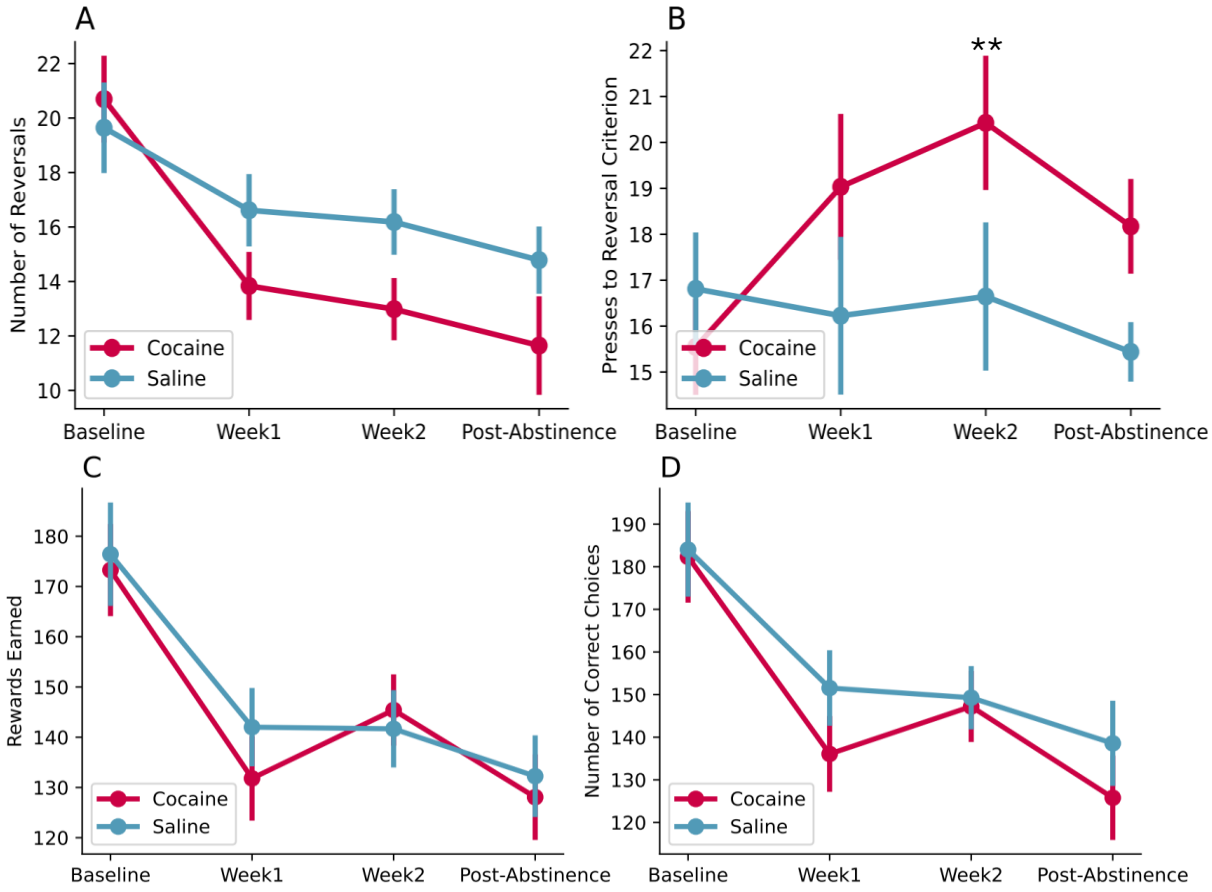
Supplementary Figure 2. **Model Validation A.** Example choice from mouse data and model generated data. **B.** Probability of left choice and of earning a reward for mouse data and model data generated with recovered parameter. No significant differences were detected between mouse data and model-generated data.

Chapter 1: Sex-Differences in Learning.

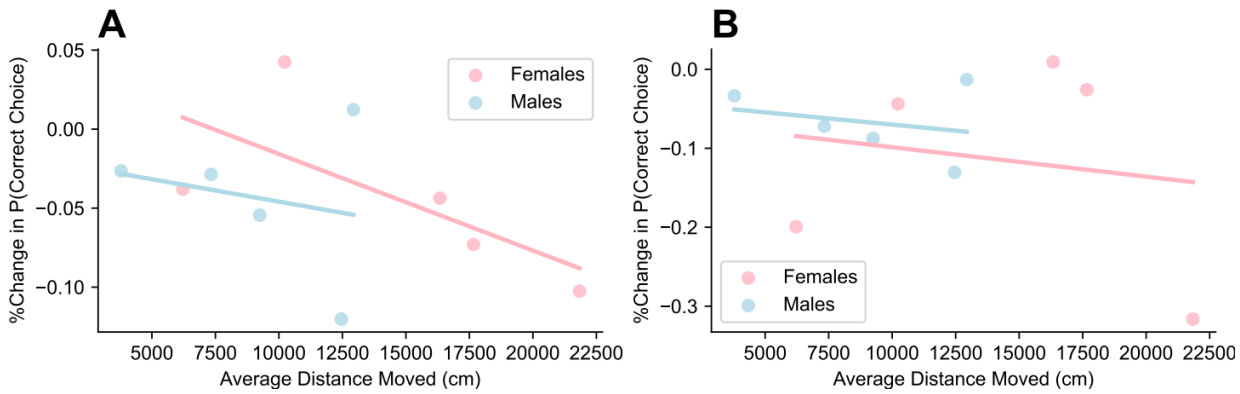


Supplementary Figure 3. **Sex-Differences during FR1 Training.** While mice displayed a preference for a lever during training, no sex-difference in side-bias were detected during **A,** 2h sessions and **B,** 30-minute sessions. The reaction time did not differ between sexes across **C,** 2h sessions and **D,** 30-minute sessions.

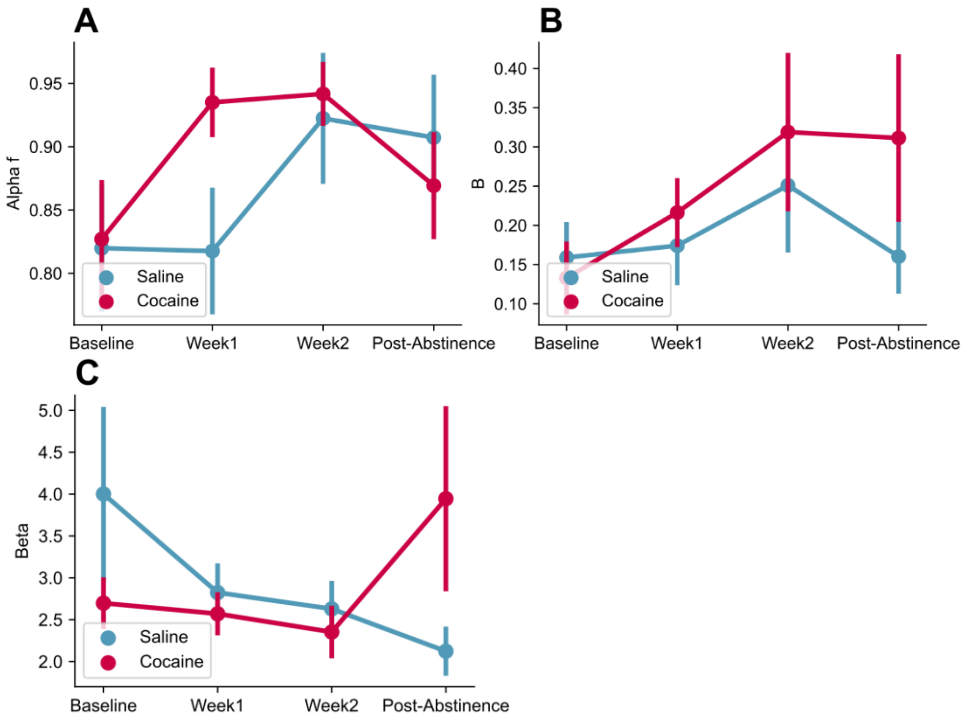
Chapter 2: Chronic Cocaine Exposure and Reinforcement Learning.



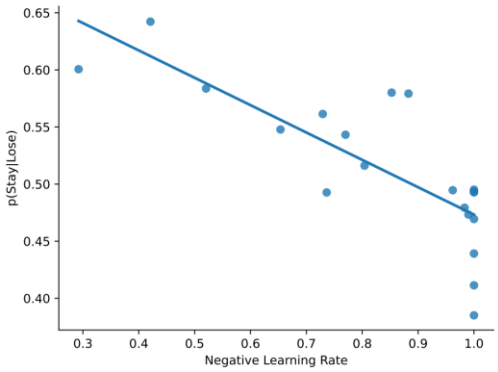
Supplementary Figure 4. Drug Differences in Other PRL Measures A. No drug significant drug differences were detected in the number of reversals achieved per session, despite a trend to significance for week 2 (Drug x Time interaction; $F(3, 414.19)=5.863$, $p<0.001$; week 2 cocaine vs saline, $t(29.6)=1.922$, $p=0.064$)**. B.** Cocaine increased the number of lever presses needed to reach the reversal criteria during week 2 only in both sexes (Drug x Time interaction; $F(3, 415.35)=3.464$, $p=0.016$; week 2 cocaine vs saline, $t(47.2)=-2.15$, $p=0.036$). **C**. The number of rewards earned per session fluctuated over time with no effect of drug nor sex (Time main effect, $F(3,30.39)=14.18$, $p<0.001$). **D**. The total number of correct choices per session fluctuated over time with no effect of drug nor sex (Time main effect, $F(3, 29.92)=9.969$, $p<0.001$).



Supplementary Figure 5. Correlation Between Decision-Making Changes and Locomotor Effects of Cocaine Injections. No significant correlations were detected between average change in choice accuracy from baseline and the average distance moved during **A,** Week 1 of injections in males (Pearson coefficient=-0.22, p=0.73) and females (Pearson coefficient=-0.70, p=0.19) and **B,** Week 2 (males, Pearson coefficient=-0.25, p=0.68; females, Pearson coefficient=-0.16, p=0.78).



Supplementary Figure 6. Cocaine Effects on RL parameters. No significant effect of time, sex nor drug were detected in **A,** forgetting rate parameter, **B,** Side bias and **C.** Inverse temperature.



Supplementary Figure 7. Linear correlation between negative learning rate values Lose-Stay ratio. The lose-stay ratio and negative learning rate displayed a significant negative correlation (Pearson correlation coefficient=-0.78; $p<0.001$)

References

Aarde, S. M., Genner, R. M., Hrncir, H., Arnold, A. P., & Jentsch, J. D. (2021). Sex chromosome complement affects multiple aspects of reversal-learning task performance in mice. Genes, Brain and Behavior, 20(1), e12685. https://doi.org/10.1111/gbb.12685

Al-Hasani, R., McCall, J. G., Shin, G., Gomez, A. M., Schmitz, G. P., Bernardi, J. M., Pyo, C.-O., Park, S. I., Marcinkiewcz, C. M., Crowley, N. A., Krashes, M. J., Lowell, B. B., Kash, T. L., Rogers, J. A., & Bruchas, M. R. (2015). Distinct Subpopulations of Nucleus Accumbens Dynorphin Neurons Drive Aversion and Reward. Neuron, 87(5), 1063–1077. https://doi.org/10.1016/j.neuron.2015.08.019

Alsiö, J., Phillips, B. U., Sala-Bayo, J., Nilsson, S. R. O., Calafat-Pla, T. C., Rizwand, A., Plumbridge, J. M., López-Cruz, L., Dalley, J. W., Cardinal, R. N., Mar, A. C., & Robbins, T. W. (2019). Dopamine D2-like receptor stimulation blocks negative feedback in visual and spatial reversal learning in the rat: Behavioural and computational evidence. Psychopharmacology, 236(8), 2307–2323. https://doi.org/10.1007/s00213-019-05296-y

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Andersen, S. L., Rutstein, M., Benzo, J. M., Hostetter, J. C., & Teicher, M. H. (1997). Sex differences in dopamine receptor overproduction and elimination: NeuroReport, 8(6), 1495–1497. https://doi.org/10.1097/00001756-199704140-00034

Anderson, L. C., & Petrovich, G. D. (2015). Renewal of conditioned responding to food cues in rats: Sex differences and relevance of estradiol. Physiology & Behavior, 151, 338–344. https://doi.org/10.1016/j.physbeh.2015.07.035

Arad, M., & Weiner, I. (2012). Abnormally rapid reversal learning and reduced response to antipsychotic drugs following ovariectomy in female rats. Psychoneuroendocrinology, 37(2), 200–212. https://doi.org/10.1016/j.psyneuen.2011.06.004

Aubele, T., & Kritzer, M. F. (2012). Androgen Influence on Prefrontal Dopamine Systems in Adult Male Rats: Localization of Cognate Intracellular Receptors in Medial Prefrontal Projections to the Ventral Tegmental Area and Effects of Gonadectomy and Hormone Replacement on Glutamate-Stimulated Extracellular Dopamine Level. Cerebral Cortex, 22(8), 1799–1812. https://doi.org/10.1093/cercor/bhr258

Bağci, B., Düsmez, S., Zorlu, N., Bahtiyar, G., Isikli, S., Bayrakci, A., Heinz, A., Schad, D. J., & Sebold, M. (2022). Computational analysis of probabilistic reversal learning deficits in male subjects with alcohol use disorder. Frontiers in Psychiatry, 13, 960238. https://doi.org/10.3389/fpsyt.2022.960238

Baker, T. E., Zeighami, Y., Dagher, A., & Holroyd, C. B. (2018). Smoking Decisions: Altered Reinforcement Learning Signals Induced by Nicotine State. Nicotine & Tobacco Research. https://doi.org/10.1093/ntr/nty136

Balconi, M., Finocchiaro, R., & Campanella, S. (2014). Reward Sensitivity, Decisional Bias, and Metacognitive Deficits in Cocaine Drug Addiction. Journal of Addiction Medicine, 8(6), 399–406. https://doi.org/10.1097/ADM.0000000000000065

Bangasser, D. A., & Cuarenta, A. (2021). Sex differences in anxiety and depression: circuits and mechanisms. Nature reviews. Neuroscience, 22(11), 674–684. https://doi.org/10.1038/s41583-021-00513-0

Bariselli, S., Fobbs, W. C., Creed, M. C., & Kravitz, A. V. (2019). A competitive model for striatal action selection. Brain Research, 1713, 70–79. https://doi.org/10.1016/j.brainres.2018.10.009

Barker, J. M., Torregrossa, M. M., Arnold, A. P., & Taylor, J. R. (2010). Dissociation of Genetic and Hormonal Influences on Sex Differences in Alcoholism-Related Behaviors. Journal of Neuroscience, 30(27), 9140–9144. https://doi.org/10.1523/JNEUROSCI.0548-10.2010

Barry, D., & Petry, N. M. (2008). Predictors of decision-making on the Iowa Gambling Task: Independent effects of lifetime history of substance use disorders and performance on the Trail Making Test. Brain and Cognition, 66(3), 243–252. https://doi.org/10.1016/j.bandc.2007.09.001

Bartolo, R., & Averbeck, B. B. (2020). Prefrontal Cortex Predicts State Switches during Reversal Learning. Neuron, 106(6), 1044-1054.e4. https://doi.org/10.1016/j.neuron.2020.03.024

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi:10.18637/jss.v067.i01

Bayer, J., Bandurski, P., & Sommer, T. (2013). Differential modulation of activity related to the anticipation of monetary gains and losses across the menstrual cycle. European Journal of Neuroscience, 38(10), 3519–3526. https://doi.org/10.1111/ejn.12347

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50(1–3), 7–15. https://doi.org/10.1016/0010-0277(94)90018-3

Bechard, A. R., LaCrosse, A., Namba, M. D., Jackson, B., & Knackstedt, L. A. (2018). Impairments in reversal learning following short access to cocaine self-administration. Drug and Alcohol Dependence, 192, 239–244. https://doi.org/10.1016/j.drugalcdep.2018.08.008

Becker, J. B., & Koob, G. F. (2016). Sex Differences in Animal Models: Focus on Addiction. Pharmacological Reviews, 68(2), 242–263. https://doi.org/10.1124/pr.115.011163

Becker, J. B., McClellan, M. L., & Reed, B. G. (2017). Sex differences, gender and addiction. Journal of Neuroscience Research, 95(1–2), 136–147. https://doi.org/10.1002/jnr.23963

Becker, J. B., McClellan, M. L., & Reed, B. G. (2017). Sex differences, gender and addiction. Journal of Neuroscience Research, 95(1–2), 136–147. https://doi.org/10.1002/jnr.23963

Bisaga, A., Aharonovich, E., Cheng, W. Y., Levin, F. R., Mariani, J. J., Raby, W. N., & Nunes, E. V. (2010). A placebo-controlled trial of memantine for cocaine dependence with high-value voucher incentives during a pre-randomization lead-in period. Drug and Alcohol Dependence, 111(1–2), 97–104. https://doi.org/10.1016/j.drugalcdep.2010.04.006

Blaes, S. L., Shimp, K. G., Betzhold, S. M., Setlow, B., & Orsini, C. A. (2022). Chronic cocaine causes age-dependent increases in risky choice in both males and females. Behavioral Neuroscience, 136(3), 243–263. https://doi.org/10.1037/bne0000509

Boulougouris, V., Castañé, A., & Robbins, T. W. (2009). Dopamine D2/D3 receptor agonist quinpirole impairs spatial reversal learning in rats: Investigation of D3 receptor involvement in persistent behavior. Psychopharmacology, 202(4), 611–620. https://doi.org/10.1007/s00213-008-1341-2

Brimblecombe, K. R., & Cragg, S. J. (2017). The Striosome and Matrix Compartments of the Striatum: A Path through the Labyrinth from Neurochemistry toward Function. ACS Chemical Neuroscience, 8(2), 235–242. https://doi.org/10.1021/acschemneuro.6b00333

Brown, V. M., Zhu, L., Solway, A., Wang, J. M., McCurry, K. L., King-Casas, B., & Chiu, P. H. (2021). Reinforcement Learning Disruptions in Individuals With Depression and Sensitivity to Symptom Change Following Cognitive Behavioral Therapy. JAMA Psychiatry, 78(10), 1113. https://doi.org/10.1001/jamapsychiatry.2021.1844

Cáceda, R., Nemeroff, C. B., & Harvey, P. D. (2014). Toward an understanding of decision making in severe mental illness. The Journal of neuropsychiatry and clinical neurosciences, 26(3), 196–213. https://doi.org/10.1176/appi.neuropsych.12110268

Calipari, E. S., Juarez, B., Morel, C., Walker, D. M., Cahill, M. E., Ribeiro, E., Roman-Ortiz, C., Ramakrishnan, C., Deisseroth, K., Han, M.-H., & Nestler, E. J. (2017). Dopaminergic dynamics underlying sex-specific cocaine reward. Nature Communications, 8(1), 13877. https://doi.org/10.1038/ncomms13877

Calu, D. J., Stalnaker, T. A., Franz, T. M., Singh, T., Shaham, Y., & Schoenbaum, G. (2007). Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. Learning & Memory, 14(5), 325–328. https://doi.org/10.1101/lm.534807

Carroll, K. M., Rounsaville, B. J., & Gawin, F. H. (1991). A Comparative Trial of Psychotherapies for Ambulatory Cocaine Abusers: Relapse Prevention and Interpersonal Psychotherapy. The American Journal of Drug and Alcohol Abuse, 17(3), 229–247. https://doi.org/10.3109/00952999109027549

Chapp, A. D., Nwakama, C. A., Jagtap, P. P., Phan, C.-M. H., Thomas, M. J., & Mermelstein, P. G. (2024). Fundamental Sex Differences in Cocaine-Induced Plasticity of Dopamine D1 Receptor– and D2 Receptor–Expressing Medium Spiny Neurons in the Mouse Nucleus Accumbens Shell. Biological Psychiatry Global Open Science, 4(3), 100295. https://doi.org/10.1016/j.bpsgos.2024.100295

Chen, C. S., Ebitz, R. B., Bindas, S. R., Redish, A. D., Hayden, B. Y., & Grissom, N. M. (2021). Divergent Strategies for Learning in Males and Females. Current Biology, 31(1), 39-50.e4. https://doi.org/10.1016/j.cub.2020.09.075

Chen, C. S., Knep, E., Han, A., Ebitz, R. B., & Grissom, N. M. (2021). Sex differences in learning from exploration. eLife, 10, e69748. https://doi.org/10.7554/eLife.69748

Chowdhury, T. G., Wallin-Miller, K. G., Rear, A. A., Park, J., Diaz, V., Simon, N. W., & Moghaddam, B. (2019). Sex differences in reward- and punishment-guided actions. Cognitive, Affective, & Behavioral Neuroscience, 19(6), 1404–1417. https://doi.org/10.3758/s13415-019-00736-w

Cocker, P. J., Rotge, J., Daniel, M., Belin-Rauscent, A., & Belin, D. (2020). Impaired decision making following escalation of cocaine self-administration predicts vulnerability to relapse in rats. Addiction Biology, 25(3), e12738. https://doi.org/10.1111/adb.12738

Cooper, S. E., Goings, S. P., Kim, J. Y., & Wood, R. I. (2014). Testosterone enhances risk tolerance without altering motor impulsivity in male rats. Psychoneuroendocrinology, 40, 201–212. https://doi.org/10.1016/j.psyneuen.2013.11.017

Cox, S. M. L., Frank, M. J., Larcher, K., Fellows, L. K., Clark, C. A., Leyton, M., & Dagher, A. (2015). Striatal D1 and D2 signaling differentially predict learning from positive and negative outcomes. NeuroImage, 109, 95–101. https://doi.org/10.1016/j.neuroimage.2014.12.070

Daniel, J. (2004). Estrogen replacement in ovariectomized rats affects strategy selection in the Morris water maze. Neurobiology of Learning and Memory, 82(2), 142–149. https://doi.org/10.1016/j.nlm.2004.06.001

Daniel, J. M., Hulst, J. L., & Berbling, J. L. (2006). Estradiol Replacement Enhances Working Memory in Middle-Aged Rats When Initiated Immediately after Ovariectomy But Not after a Long-Term Period of Ovarian Hormone Deprivation. Endocrinology, 147(1), 607–614. https://doi.org/10.1210/en.2005-0998

Dazzi, L., Seu, E., Cherchi, G., Barbieri, P. P., Matzeu, A., & Biggio, G. (2007). Estrous Cycle-Dependent Changes in Basal and Ethanol-Induced Activity of Cortical Dopaminergic Neurons in the Rat. Neuropsychopharmacology, 32(4), 892–901. https://doi.org/10.1038/sj.npp.1301150

Degenhardt, L., Charlson, F., Ferrari, A., Santomauro, D., Erskine, H., Mantilla-Herrara, A., Whiteford, H., Leung, J., Naghavi, M., Griswold, M., Rehm, J., Hall, W., Sartorius, B., Scott, J., Vollset, S. E., Knudsen, A. K., Haro, J. M., Patton, G., Kopec, J., … Vos, T. (2018). The global burden of disease attributable to alcohol and drug use in 195 countries

and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. The Lancet Psychiatry, 5(12), 987–1012. https://doi.org/10.1016/S2215-0366(18)30337-7

Dreher, J.-C., Schmidt, P. J., Kohn, P., Furman, D., Rubinow, D., & Berman, K. F. (2007). Menstrual cycle phase modulates reward-related neural function in women. Proceedings of the National Academy of Sciences, 104(7), 2465–2470. https://doi.org/10.1073/pnas.0605569104

Dutra, L., Stathopoulou, G., Basden, S. L., Leyro, T. M., Powers, M. B., & Otto, M. W. (2008). A Meta-Analytic Review of Psychosocial Interventions for Substance Use Disorders. American Journal of Psychiatry, 165(2), 179–187. https://doi.org/10.1176/appi.ajp.2007.06111851

Ersche, K. D., Gillan, C. M., Jones, P. S., Williams, G. B., Ward, L. H. E., Luijten, M., De Wit, S., Sahakian, B. J., Bullmore, E. T., & Robbins, T. W. (2016). Carrots and sticks fail to change behavior in cocaine addiction. Science, 352(6292), 1468–1471. https://doi.org/10.1126/science.aaf3700

Ersche, K. D., Roiser, J. P., Abbott, S., Craig, K. J., Müller, U., Suckling, J., Ooi, C., Shabbir, S. S., Clark, L., Sahakian, B. J., Fineberg, N. A., Merlo-Pich, E. V., Robbins, T. W., & Bullmore, E. T. (2011). Response Perseveration in Stimulant Dependence Is Associated with Striatal Dysfunction and Can Be Ameliorated by a D2/3 Receptor Agonist. Biological Psychiatry, 70(8), 754–762. https://doi.org/10.1016/j.biopsych.2011.06.033

Ersche, K. D., Roiser, J. P., Robbins, T. W., & Sahakian, B. J. (2008). Chronic cocaine but not chronic amphetamine use is associated with perseverative responding in humans. Psychopharmacology, 197(3), 421–431. https://doi.org/10.1007/s00213-007-1051-1

Evans, K. L., & Hampson, E. (2015). Sex-dependent effects on tasks assessing reinforcement learning and interference inhibition. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.01044

Everitt, B. J., Belin, D., Economidou, D., Pelloux, Y., Dalley, J. W., & Robbins, T. W. (2008). Neural mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1507), 3125–3135. https://doi.org/10.1098/rstb.2008.0089

Fellows, L. K. (2004). Different Underlying Impairments in Decision-making Following Ventromedial and Dorsolateral Frontal Lobe Damage in Humans. Cerebral Cortex, 15(1), 58–63. https://doi.org/10.1093/cercor/bhh108

Ferland, J. N., & Winstanley, C. A. (2017). Risk-preferring rats make worse decisions and show increased incubation of craving after cocaine self-administration. Addiction Biology, 22(4), 991–1001. https://doi.org/10.1111/adb.12388

Fernández-Serrano, M. J., Perales, J. C., Moreno-López, L., Pérez-García, M., & Verdejo-García, A. (2012). Neuropsychological profiling of impulsivity and compulsivity in cocaine dependent individuals. Psychopharmacology, 219(2), 673–683. https://doi.org/10.1007/s00213-011-2485-z

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning.

Proceedings of the National Academy of Sciences, 104(41), 16311–16316. https://doi.org/10.1073/pnas.0706111104

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. Science, 306(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Franken, I. H. A., Van Strien, J. W., Nijs, I., & Muris, P. (2008). Impulsivity is associated with behavioral decision-making deficits. Psychiatry Research, 158(2), 155–163. https://doi.org/10.1016/j.psychres.2007.06.002

Georgiou, P., Zanos, P., Bhat, S., Tracy, J. K., Merchenthaler, I. J., McCarthy, M. M., & Gould, T. D. (2018). Dopamine and Stress System Modulation of Sex Differences in Decision Making. Neuropsychopharmacology, 43(2), 313–324. https://doi.org/10.1038/npp.2017.161

Gerfen, C. R., & Surmeier, D. J. (2011). Modulation of Striatal Projection Systems by Dopamine. Annual Review of Neuroscience, 34(1), 441–466. https://doi.org/10.1146/annurev-neuro-061010-113641

Gowin, J. L., Sloan, M. E., Ramchandani, V. A., Paulus, M. P., & Lane, S. D. (2018). Differences in decision-making as a function of drug of choice. Pharmacology Biochemistry and Behavior, 164, 118–124. https://doi.org/10.1016/j.pbb.2017.09.007

Groman, S. M., Hillmer, A. T., Liu, H., Fowles, K., Holden, D., Morris, E. D., Lee, D., & Taylor, J. R. (2020a). Dysregulation of Decision Making Related to Metabotropic Glutamate 5, but Not Midbrain D3, Receptor Availability Following Cocaine Self-administration in

Rats. Biological Psychiatry, 88(10), 777–787. https://doi.org/10.1016/j.biopsych.2020.06.020

Groman, S. M., Hillmer, A. T., Liu, H., Fowles, K., Holden, D., Morris, E. D., Lee, D., & Taylor, J. R. (2020b). Midbrain D3 Receptor Availability Predicts Escalation in Cocaine Self-administration. Biological Psychiatry, 88(10), 767–776. https://doi.org/10.1016/j.biopsych.2020.02.017

Haluk, D. M., & Floresco, S. B. (2009). Ventral Striatal Dopamine Modulation of Different Forms of Behavioral Flexibility. Neuropsychopharmacology, 34(8), 2041–2052. https://doi.org/10.1038/npp.2009.21

Hasbi, A., Nguyen, T., Rahal, H., Manduca, J. D., Miksys, S., Tyndale, R. F., Madras, B. K., Perreault, M. L., & George, S. R. (2020). Sex difference in dopamine D1-D2 receptor complex expression and signaling affects depression- and anxiety-like behaviors. Biology of Sex Differences, 11(1), 8. https://doi.org/10.1186/s13293-020-00285-9

Heywood, L. H. (1980). Testosterone Levels in the Male Laboratory Rat: Variation under Experimental Conditions. International Journal of Andrology, 3(1–6), 519–529. https://doi.org/10.1111/j.1365-2605.1980.tb00140.x

Himanshu Gangal, Xueyi Xie, Zhenbo Huang, Yifeng Cheng, Xuehua Wang, Jiayi Lu, Xiaowen Zhuang, Essoh, A., Yufei Huang, Ruifeng Chen, Smith, L. N., Smith, R. J., & Wang, J. (2023). Drug Reinforcement Impairs Cognitive Flexibility by Inhibiting Striatal Cholinergic Neurons. https://doi.org/10.5281/ZENODO.7948766

Imperato, A., & Di Chiara, G. (1985). Dopamine release and metabolism in awake rats after systemic neuroleptics as studied by trans-striatal dialysis. The Journal of Neuroscience, 5(2), 297–306. https://doi.org/10.1523/JNEUROSCI.05-02-00297.1985

Islas-Preciado, D., Wainwright, S. R., Sniegocki, J., Lieblich, S. E., Yagi, S., Floresco, S. B., & Galea, L. A. M. (2020). Risk-based decision making in rats: Modulation by sex and amphetamine. Hormones and Behavior, 125, 104815. https://doi.org/10.1016/j.yhbeh.2020.104815

Islas-Preciado, D., Wainwright, S. R., Sniegocki, J., Lieblich, S. E., Yagi, S., Floresco, S. B., & Galea, L. A. M. (2020). Risk-based decision making in rats: Modulation by sex and amphetamine. Hormones and Behavior, 125, 104815. https://doi.org/10.1016/j.yhbeh.2020.104815

Iyer, E. S., Kairiss, M. A., Liu, A., Otto, A. R., & Bagot, R. C. (2020). Probing relationships between reinforcement learning and simple behavioral strategies to understand probabilistic reward learning. *Journal of neuroscience methods*, *341*, 108777. https://doi.org/10.1016/j.jneumeth.2020.108777

Izquierdo, A., & Jentsch, J. D. (2012). Reversal learning as a measure of impulsive and compulsive behavior in addictions. Psychopharmacology, 219(2), 607–620. https://doi.org/10.1007/s00213-011-2579-7

Jacome, L. F., Gautreaux, C., Inagaki, T., Mohan, G., Alves, S., Lubbers, L. S., & Luine, V. (2010). Estradiol and ERβ agonists enhance recognition memory, and DPN, an ERβ agonist, alters brain monoamines. Neurobiology of Learning and Memory, 94(4), 488–498. https://doi.org/10.1016/j.nlm.2010.08.016

Jenni, N. L., Li, Y. T., & Floresco, S. B. (2021). Medial orbitofrontal cortex dopamine D1/D2 receptors differentially modulate distinct forms of probabilistic decision-making. Neuropsychopharmacology, 46(7), 1240–1251. https://doi.org/10.1038/s41386-020-00931-1

Kampman, K. M. (2019). The treatment of cocaine use disorder. Science Advances, 5(10), eaax1532. https://doi.org/10.1126/sciadv.aax1532

Kanen, J. W., Ersche, K. D., Fineberg, N. A., Robbins, T. W., & Cardinal, R. N. (2019). Computational modelling reveals contrasting effects on reinforcement learning and cognitive flexibility in stimulant use disorder and obsessive-compulsive disorder: Remediating effects of dopaminergic D2/3 receptor agents. Psychopharmacology, 236(8), 2337–2358. https://doi.org/10.1007/s00213-019-05325-w

Kim, J., Park, B.-H., Lee, J. H., Park, S. K., & Kim, J.-H. (2011). Cell Type-Specific Alterations in the Nucleus Accumbens by Repeated Exposures to Cocaine. Biological Psychiatry, 69(11), 1026–1034. https://doi.org/10.1016/j.biopsych.2011.01.013

Kjome, K. L., Lane, S. D., Schmitz, J. M., Green, C., Ma, L., Prasla, I., Swann, A. C., & Moeller, F. G. (2010). Relationship between impulsivity and decision making in cocaine dependence. Psychiatry Research, 178(2), 299–304. https://doi.org/10.1016/j.psychres.2009.11.024

Kokras, N., Pastromas, N., Papasava, D., De Bournonville, C., Cornil, C. A., & Dalla, C. (2018). Sex differences in behavioral and neurochemical effects of gonadectomy and aromatase inhibition in rats. Psychoneuroendocrinology, 87, 93–107. https://doi.org/10.1016/j.psyneuen.2017.10.007

Kritzer, M. F., & Creutz, L. M. (2008). Region and Sex Differences in Constituent Dopamine Neurons and Immunoreactivity for Intracellular Estrogen and Androgen Receptors in Mesocortical Projections in Rats. The Journal of Neuroscience, 28(38), 9525–9535. https://doi.org/10.1523/JNEUROSCI.2637-08.2008

Krueger, D. D., Howell, J. L., Oo, H., Olausson, P., Taylor, J. R., & Nairn, A. C. (2009). Prior chronic cocaine exposure in mice induces persistent alterations in cognitive function. Behavioural Pharmacology, 20(8), 695–704. https://doi.org/10.1097/FBP.0b013e328333a2bb

Kupchik, Y. M., Brown, R. M., Heinsbroek, J. A., Lobo, M. K., Schwartz, D. J., & Kalivas, P. W. (2015). Coding the direct/indirect pathways by D1 and D2 receptors is not valid for accumbens projections. Nature Neuroscience, 18(9), 1230–1232. https://doi.org/10.1038/nn.4068

Kutlu, M. G., Zachry, J. E., Brady, L. J., Melugin, P. R., Kelly, S. J., Sanders, C., Tat, J., Johnson, A. R., Thibeault, K., Lopez, A. J., Siciliano, C. A., & Calipari, E. S. (2020). A novel multidimensional reinforcement task in mice elucidates sex-specific behavioral strategies. Neuropsychopharmacology, 45(9), 1463–1472. https://doi.org/10.1038/s41386-020-0692-1

LeBlanc, K. H., Maidment, N. T., & Ostlund, S. B. (2013). Repeated Cocaine Exposure Facilitates the Expression of Incentive Motivation and Induces Habitual Control in Rats. PLoS ONE, 8(4), e61355. https://doi.org/10.1371/journal.pone.0061355

Lefner, M. J., Dejeux, M. I., & Wanat, M. J. (2022). Sex Differences in Behavioral Responding and Dopamine Release during Pavlovian Learning. Eneuro, 9(2), ENEURO.0050-22.2022. https://doi.org/10.1523/ENEURO.0050-22.2022

Liley, A. E., Gabriel, D. B. K., Sable, H. J., & Simon, N. W. (2019). Sex Differences and Effects of Predictive Cues on Delayed Punishment Discounting. Eneuro, 6(4), ENEURO.0225-19.2019. https://doi.org/10.1523/ENEURO.0225-19.2019

Liley, A. E., Gabriel, D. B. K., Sable, H. J., & Simon, N. W. (2019). Sex Differences and Effects of Predictive Cues on Delayed Punishment Discounting. Eneuro, 6(4), ENEURO.0225-19.2019. https://doi.org/10.1523/ENEURO.0225-19.2019

Lim, T. V., Cardinal, R. N., Bullmore, E. T., Robbins, T. W., & Ersche, K. D. (2021). Impaired Learning From Negative Feedback in Stimulant Use Disorder: Dopaminergic Modulation. International Journal of Neuropsychopharmacology, 24(11), 867–878. https://doi.org/10.1093/ijnp/pyab041

Lim, T. V., Cardinal, R. N., Bullmore, E. T., Robbins, T. W., & Ersche, K. D. (2021). Impaired Learning From Negative Feedback in Stimulant Use Disorder: Dopaminergic Modulation. International Journal of Neuropsychopharmacology, 24(11), 867–878. https://doi.org/10.1093/ijnp/pyab041

Liu, C. L., Wang, Y. K., Jin, G. Z., Shi, W. X., & Gao, M. (2018). Cocaine-induced locomotor sensitization associates with slow oscillatory firing of neurons in the ventral tegmental area. *Scientific reports*, *8*(1), 3274. https://doi.org/10.1038/s41598-018-21592-7

Lui, J. H., Nguyen, N. D., Grutzner, S. M., Darmanis, S., Peixoto, D., Wagner, M. J., Allen, W. E., Kebschull, J. M., Richman, E. B., Ren, J., Newsome, W. T., Quake, S. R., & Luo, L.

(2021). Differential encoding in prefrontal cortex projection neuron classes across cognitive tasks. Cell, 184(2), 489-506.e26. https://doi.org/10.1016/j.cell.2020.11.046

Macpherson, T., Kim, J. Y., & Hikida, T. (2022). Nucleus Accumbens Core Dopamine D2 Receptor-Expressing Neurons Control Reversal Learning but Not Set-Shifting in Behavioral Flexibility in Male Mice. Frontiers in Neuroscience, 16, 885380. https://doi.org/10.3389/fnins.2022.885380

Mar, A. C., Walker, A. L. J., Theobald, D. E., Eagle, D. M., & Robbins, T. W. (2011). Dissociable Effects of Lesions to Orbitofrontal Cortex Subregions on Impulsive Choice in the Rat. The Journal of Neuroscience, 31(17), 6398–6404. https://doi.org/10.1523/JNEUROSCI.6620-10.2011

Marcellino, D., Kehr, J., Agnati, L. F., & Fuxe, K. (2012). Increased affinity of dopamine for D 2 -like versus D 1 -like receptors. Relevance for volume transmission in interpreting PET findings. Synapse, 66(3), 196–203. https://doi.org/10.1002/syn.21501

Maude-Griffin, P. M., Hohenstein, J. M., Humfleet, G. L., Reilly, P. M., Tusel, D. J., & Hall, S. M. (1998). Superior efficacy of cognitive-behavioral therapy for urban crack cocaine abusers: Main and matching effects. Journal of Consulting and Clinical Psychology, 66(5), 832–837. https://doi.org/10.1037/0022-006X.66.5.832

McCarthy, M. M., & Arnold, A. P. (2011). Reframing sexual differentiation of the brain. Nature Neuroscience, 14(6), 677–683. https://doi.org/10.1038/nn.2834

McCracken, C. B., & Grace, A. A. (2013). Persistent Cocaine-Induced Reversal Learning Deficits Are Associated with Altered Limbic Cortico-Striatal Local Field Potential

Synchronization. The Journal of Neuroscience, 33(44), 17469–17482. https://doi.org/10.1523/JNEUROSCI.1440-13.2013

Mitchell, M. R., Weiss, V. G., Beas, B. S., Morgan, D., Bizon, J. L., & Setlow, B. (2014). Adolescent Risk Taking, Cocaine Self-Administration, and Striatal Dopamine Signaling. Neuropsychopharmacology, 39(4), 955–962. https://doi.org/10.1038/npp.2013.295

Myers, C. E., Sheynin, J., Balsdon, T., Luzardo, A., Beck, K. D., Hogarth, L., Haber, P., & Moustafa, A. A. (2016). Probabilistic reward- and punishment-based learning in opioid addiction: Experimental and computational data. Behavioural Brain Research, 296, 240–248. https://doi.org/10.1016/j.bbr.2015.09.018

Nader, M. A., & Czoty, P. W. (2005). PET Imaging of Dopamine D2 Receptors in Monkey Models of Cocaine Abuse: Genetic Predisposition Versus Environmental Modulation. American Journal of Psychiatry, 162(8), 1473–1482. https://doi.org/10.1176/appi.ajp.162.8.1473

Nejtek, V. A., Kaiser, K. A., Zhang, B., & Djokovic, M. (2013). Iowa Gambling Task scores predict future drug use in bipolar disorder outpatients with stimulant dependence. Psychiatry Research, 210(3), 871–879. https://doi.org/10.1016/j.psychres.2013.08.021

Nigro, S. E., Wu, M., C. Juliano, A., Flynn, B., Lu, L. H., Landay, A. L., French, A. L., & Yang, S. (2021). Effects of cocaine and HIV on decision-making abilities. Journal of NeuroVirology, 27(3), 422–433. https://doi.org/10.1007/s13365-021-00965-1

Nunamaker, E. A., & Turner, P. V. (2023). Unmasking the Adverse Impacts of Sex Bias on Science and Research Animal Welfare. Animals, 13(17), 2792. https://doi.org/10.3390/ani13172792

Orendain-Jaime, E. N., Ortega-Ibarra, J. M., & López-Pérez, S. J. (2016). Evidence of sexual dimorphism in D1 and D2 dopaminergic receptors expression in frontal cortex and striatum of young rats. Neurochemistry International, 100, 62–66. https://doi.org/10.1016/j.neuint.2016.09.001

Orsini, C. A., Blaes, S. L., Dragone, R. J., Betzhold, S. M., Finner, A. M., Bizon, J. L., & Setlow, B. (2020). Distinct relationships between risky decision making and cocaine self-administration under short- and long-access conditions. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 98, 109791. https://doi.org/10.1016/j.pnpbp.2019.109791

Orsini, C. A., Blaes, S. L., Hernandez, C. M., Betzhold, S. M., Perera, H., Wheeler, A.-R., Ten Eyck, T. W., Garman, T. S., Bizon, J. L., & Setlow, B. (2021). Regulation of risky decision making by gonadal hormones in males and females. Neuropsychopharmacology, 46(3), 603–613. https://doi.org/10.1038/s41386-020-00827-0

Orsini, C. A., Heshmati, S. C., Garman, T. S., Wall, S. C., Bizon, J. L., & Setlow, B. (2018). Contributions of medial prefrontal cortex to decision making involving risk of punishment. Neuropharmacology, 139, 205–216. https://doi.org/10.1016/j.neuropharm.2018.07.018

Orsini, C. A., Trotta, R. T., Bizon, J. L., & Setlow, B. (2015). Dissociable Roles for the Basolateral Amygdala and Orbitofrontal Cortex in Decision-Making under Risk of Punishment. The Journal of Neuroscience, 35(4), 1368–1379. https://doi.org/10.1523/JNEUROSCI.3586-14.2015

Orsini, C. A., Truckenbrod, L. M., & Wheeler, A.-R. (2022). Regulation of sex differences in risk-based decision making by gonadal hormones: Insights from rodent models. Behavioural Processes, 200, 104663. https://doi.org/10.1016/j.beproc.2022.104663

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

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

Quinn, J. J., Hitchcott, P. K., Umeda, E. A., Arnold, A. P., & Taylor, J. R. (2007). Sex chromosome complement regulates habit formation. Nature Neuroscience, 10(11), 1398–1400. https://doi.org/10.1038/nn1994

Rai, L. A., O'Halloran, L., Jollans, L., Vahey, N., O'Brolchain, C., & Whelan, R. (2019). Individual differences in learning from probabilistic reward and punishment predicts smoking status. Addictive Behaviors, 88, 73–76. https://doi.org/10.1016/j.addbeh.2018.08.019

Rohsenow, D. J., Monti, P. M., Martin, R. A., Michalec, E., & Abrams, D. B. (2000). Brief coping skills treatment for cocaine abuse: 12-month substance use outcomes. Journal of Consulting and Clinical Psychology, 68(3), 515–520. https://doi.org/10.1037/0022-006X.68.3.515

Romero Pinto, S., & Uchida, N. (2023). Tonic dopamine and biases in value learning linked through a biologically inspired reinforcement learning model. https://doi.org/10.1101/2023.11.10.566580

Sárvári, M., Deli, L., Kocsis, P., Márk, L., Maász, G., Hrabovszky, E., Kalló, I., Gajári, D., Vastagh, C., Sümegi, B., Tihanyi, K., & Liposits, Z. (2014). Estradiol and isotype-selective estrogen receptor agonists modulate the mesocortical dopaminergic system in gonadectomized female rats. Brain Research, 1583, 1–11. https://doi.org/10.1016/j.brainres.2014.06.020

Savell, K. E., Tuscher, J. J., Zipperly, M. E., Duke, C. G., Phillips, R. A., Bauman, A. J., Thukral, S., Sultan, F. A., Goska, N. A., Ianov, L., & Day, J. J. (2020). A dopamine-induced gene expression signature regulates neuronal function and cocaine response. Science Advances, 6(26), eaba4221. https://doi.org/10.1126/sciadv.aba4221

Schmitz, J. M., Mooney, M. E., Green, C. E., Lane, S. D., Steinberg, J. L., Swann, A. C., & Moeller, F. G. (2009). Baseline Neurocognitive Profiles Differentiate Abstainers and Non-Abstainers in a Cocaine Clinical Trial. Journal of Addictive Diseases, 28(3), 250–257. https://doi.org/10.1080/10550880903028502

Schoenbaum, G., & Setlow, B. (2005). Cocaine Makes Actions Insensitive to Outcomes but not Extinction: Implications for Altered Orbitofrontal–Amygdalar Function. Cerebral Cortex, 15(8), 1162–1169. https://doi.org/10.1093/cercor/bhh216

Schoenbaum, G., Saddoris, M. P., Ramus, S. J., Shaham, Y., & Setlow, B. (2004). Cocaine-experienced rats exhibit learning deficits in a task sensitive to orbitofrontal cortex lesions.

European Journal of Neuroscience, 19(7), 1997–2002. https://doi.org/10.1111/j.1460-9568.2004.03274.x

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. Neuron, 91(6), 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. The Journal of Neuroscience, 13(3), 900–913. https://doi.org/10.1523/JNEUROSCI.13-03-00900.1993

Seedat, S., Scott, K. M., Angermeyer, M. C., Berglund, P., Bromet, E. J., Brugha, T. S., Demyttenaere, K., De Girolamo, G., Haro, J. M., Jin, R., Karam, E. G., Kovess-Masfety, V., Levinson, D., Medina Mora, M. E., Ono, Y., Ormel, J., Pennell, B.-E., Posada-Villa, J., Sampson, N. A., … Kessler, R. C. (2009). Cross-National Associations Between Gender and Mental Disorders in the World Health Organization World Mental Health Surveys. Archives of General Psychiatry, 66(7), 785. https://doi.org/10.1001/archgenpsychiatry.2009.36

Seu, E., Groman, S. M., Arnold, A. P., & Jentsch, J. D. (2014). Sex chromosome complement influences operant responding for a palatable food in mice. Genes, Brain and Behavior, 13(6), 527–534. https://doi.org/10.1111/gbb.12143

Simon, N. W., Montgomery, K. S., Beas, B. S., Mitchell, M. R., LaSarge, C. L., Mendez, I. A., Bañuelos, C., Vokes, C. M., Taylor, A. B., Haberman, R. P., Bizon, J. L., & Setlow, B.

(2011). Dopaminergic Modulation of Risky Decision-Making. The Journal of Neuroscience, 31(48), 17460–17470. https://doi.org/10.1523/JNEUROSCI.3772-11.2011

Smith, M. S., Freeman, M. E., & Neill, J. D. (1975). The Control of Progesterone Secretion During the Estrous Cycle and Early Pseudopregnancy in the Rat: Prolactin, Gonadotropin and Steroid Levels Associated with Rescue of the Corpus Luteum of Pseudopregnancy 1 2. Endocrinology, 96(1), 219–226. https://doi.org/10.1210/endo-96-1-219

Soares-Cunha, C., De Vasconcelos, N. A. P., Coimbra, B., Domingues, A. V., Silva, J. M., Loureiro-Campos, E., Gaspar, R., Sotiropoulos, I., Sousa, N., & Rodrigues, A. J. (2020). Nucleus accumbens medium spiny neurons subtypes signal both reward and aversion. Molecular Psychiatry, 25(12), 3241–3255. https://doi.org/10.1038/s41380-019-0484-3

St. Onge, J. R., & Floresco, S. B. (2010). Prefrontal Cortical Contribution to Risk-Based Decision Making. Cerebral Cortex, 20(8), 1816–1828. https://doi.org/10.1093/cercor/bhp250

St. Onge, J. R., Abhari, H., & Floresco, S. B. (2011). Dissociable Contributions by Prefrontal D1 and D2 Receptors to Risk-Based Decision Making. Journal of Neuroscience, 31(23), 8625–8633. https://doi.org/10.1523/JNEUROSCI.1020-11.2011

Stevens, L., Betanzos-Espinosa, P., Crunelle, C. L., Vergara-Moragues, E., Roeyers, H., Lozano, O., Dom, G., Gonzalez-Saiz, F., Vanderplasschen, W., Verdejo-García, A., & Pérez-García, M. (2013). Disadvantageous Decision-Making as a Predictor of Drop-Out among Cocaine-Dependent Individuals in Long-Term Residential Treatment. Frontiers in Psychiatry, 4. https://doi.org/10.3389/fpsyt.2013.00149

Stopper, C. M., Green, E. B., & Floresco, S. B. (2014). Selective Involvement by the Medial Orbitofrontal Cortex in Biasing Risky, But Not Impulsive, Choice. Cerebral Cortex, 24(1), 154–162. https://doi.org/10.1093/cercor/bhs297

Taylor, S., Lavalley, C. A., Hakimi, N., Stewart, J. L., Ironside, M., Zheng, H., White, E., Guinjoan, S., Paulus, M. P., & Smith, R. (2023). Active learning impairments in substance use disorders when resolving the explore-exploit dilemma: A replication and extension of previous computational modeling results. Drug and Alcohol Dependence, 252, 110945. https://doi.org/10.1016/j.drugalcdep.2023.110945

Tepper, J. M., Tecuapetla, F., Koós, T., & Ibáñez-Sandoval, O. (2010). Heterogeneity and Diversity of Striatal GABAergic Interneurons. Frontiers in Neuroanatomy, 4. https://doi.org/10.3389/fnana.2010.00150

Thompson, L. L., Claus, E. D., Mikulich-Gilbertson, S. K., Banich, M. T., Crowley, T., Krmpotich, T., Miller, D., & Tanabe, J. (2012). Negative reinforcement learning is affected in substance dependence. Drug and Alcohol Dependence, 123(1–3), 84–90. https://doi.org/10.1016/j.drugalcdep.2011.10.017

Truckenbrod, L. M., Betzhold, S. M., Wheeler, A.-R., Shallcross, J., Singhal, S., Harden, S., Schwendt, M., Frazier, C. J., Bizon, J. L., Setlow, B., & Orsini, C. A. (2023). Circuit and Cell-Specific Contributions to Decision Making Involving Risk of Explicit Punishment in Male and Female Rats. The Journal of Neuroscience, 43(26), 4837–4855. https://doi.org/10.1523/JNEUROSCI.0276-23.2023

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, 238, 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., 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 Der Plas, E. A. A., Crone, E. A., Van Den Wildenberg, W. P. M., Tranel, D., & Bechara, A. (2009). Executive control deficits in substance-dependent individuals: A comparison of alcohol, cocaine, and methamphetamine and of men and women. Journal of Clinical and Experimental Neuropsychology, 31(6), 706–719. https://doi.org/10.1080/13803390802484797

Verdejo-Garcia, A., Albein-Urios, N., Martinez-Gonzalez, J. M., Civit, E., De La Torre, R., & Lozano, O. (2014). Decision-making impairment predicts 3-month hair-indexed cocaine relapse. Psychopharmacology, 231(21), 4179–4187. https://doi.org/10.1007/s00213-014-3563-9

Verdejo-Garcia, A., Benbrook, A., Funderburk, F., David, P., Cadet, J.-L., & Bolla, K. I. (2007). The differential relationship between cocaine use and marijuana use on decision-making

performance over repeat testing with the Iowa Gambling Task. Drug and Alcohol Dependence, 90(1), 2–11. https://doi.org/10.1016/j.drugalcdep.2007.02.004

Verdejo-Garcia, A., Clark, L., Verdejo-Román, J., Albein-Urios, N., Martinez-Gonzalez, J. M., Gutierrez, B., & Soriano-Mas, C. (2015a). Neural substrates of cognitive flexibility in cocaine and gambling addictions. British Journal of Psychiatry, 207(2), 158–164. https://doi.org/10.1192/bjp.bp.114.152223

Verdejo-Garcia, A., Clark, L., Verdejo-Román, J., Albein-Urios, N., Martinez-Gonzalez, J. M., Gutierrez, B., & Soriano-Mas, C. (2015b). Neural substrates of cognitive flexibility in cocaine and gambling addictions. British Journal of Psychiatry, 207(2), 158–164. https://doi.org/10.1192/bjp.bp.114.152223

Verdejo-Garcia, A., Garcia-Fernandez, G., & Dom, G. (2019). Cognition and addiction. Dialogues in Clinical Neuroscience, 21(3), 281–290. https://doi.org/10.31887/DCNS.2019.21.3/gdom

Verharen, J. P. H., De Jong, J. W., Roelofs, T. J. M., Huffels, C. F. M., Van Zessen, R., Luijendijk, M. C. M., Hamelink, R., Willuhn, I., Den Ouden, H. E. M., Van Der Plasse, G., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2018). A neuronal mechanism underlying decision-making deficits during hyperdopaminergic states. Nature Communications, 9(1), 731. https://doi.org/10.1038/s41467-018-03087-1

Villiamma, P., Casby, J., & Groman, S. M. (2022). Adolescent reinforcement-learning trajectories predict cocaine-taking behaviors in adult male and female rats. Psychopharmacology, 239(9), 2885–2901. https://doi.org/10.1007/s00213-022-06174-w

Volkow, N. D., & Blanco, C. (2023). Substance use disorders: A comprehensive update of classification, epidemiology, neurobiology, clinical aspects, treatment and prevention. World Psychiatry, 22(2), 203–229. https://doi.org/10.1002/wps.21073

Volkow, N. D., Chang, L., Wang, G.-J., Fowler, J. S., Ding, Y.-S., Sedler, M., Logan, J., Franceschi, D., Gatley, J., Hitzemann, R., Gifford, A., Wong, C., & Pappas, N. (2001). Low Level of Brain Dopamine D 2 Receptors in Methamphetamine Abusers: Association With Metabolism in the Orbitofrontal Cortex. American Journal of Psychiatry, 158(12), 2015–2021. https://doi.org/10.1176/appi.ajp.158.12.2015

Volkow, N. D., Fowler, J. S., Wang, G.-J., Swanson, J. M., & Telang, F. (2007). Dopamine in Drug Abuse and Addiction: Results of Imaging Studies and Treatment Implications. Archives of Neurology, 64(11), 1575. https://doi.org/10.1001/archneur.64.11.1575

Volkow, N. D., Michaelides, M., & Baler, R. (2019). The Neuroscience of Drug Reward and Addiction. *Physiological reviews*, *99*(4), 2115–2140. https://doi.org/10.1152/physrev.00014.2018

Wallin, K. G., Alves, J. M., & Wood, R. I. (2015). Anabolic–androgenic steroids and decision making: Probability and effort discounting in male rats. Psychoneuroendocrinology, 57, 84–92. https://doi.org/10.1016/j.psyneuen.2015.03.023

Wallin-Miller, K. G., Kreutz, F., Li, G., & Wood, R. I. (2018). Anabolic-androgenic steroids (AAS) increase sensitivity to uncertainty by inhibition of dopamine D1 and D2 receptors. Psychopharmacology, 235(4), 959–969. https://doi.org/10.1007/s00213-017-4810-7

Wang, J. M., Zhu, L., Brown, V. M., De La Garza, R., Newton, T., King-Casas, B., & Chiu, P. H. (2019). In Cocaine Dependence, Neural Prediction Errors During Loss Avoidance Are

Increased With Cocaine Deprivation and Predict Drug Use. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4(3), 291–299. https://doi.org/10.1016/j.bpsc.2018.07.009

Wang, J. M., Zhu, L., Brown, V. M., De La Garza, R., Newton, T., King-Casas, B., & Chiu, P. H. (2019). In Cocaine Dependence, Neural Prediction Errors During Loss Avoidance Are Increased With Cocaine Deprivation and Predict Drug Use. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4(3), 291–299. https://doi.org/10.1016/j.bpsc.2018.07.009

Waskom, M. L., (2021). seaborn: statistical data visualization. Journal of Open Source Software, 6(60), 3021, https://doi.org/10.21105/joss.03021

Wilson, R. C., & Niv, Y. (2012). Inferring Relevance in a Changing World. Frontiers in Human Neuroscience, 5. https://doi.org/10.3389/fnhum.2011.00189

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. Neuron, 81(2), 267–279. https://doi.org/10.1016/j.neuron.2013.11.005

Wissman, A. M., McCollum, A. F., Huang, G. Z., Nikrodhanond, A. A., & Woolley, C. S. (2011). Sex differences and effects of cocaine on excitatory synapses in the nucleus accumbens. *Neuropharmacology*, *61*(1-2), 217–227. https://doi.org/10.1016/j.neuropharm.2011.04.002

Yoest, K. E., Quigley, J. A., & Becker, J. B. (2018). Rapid effects of ovarian hormones in dorsal striatum and nucleus accumbens. Hormones and Behavior, 104, 119–129. https://doi.org/10.1016/j.yhbeh.2018.04.002

Zalocusky, K. A., Ramakrishnan, C., Lerner, T. N., Davidson, T. J., Knutson, B., & Deisseroth, K. (2016). Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. Nature, 531(7596), 642–646. https://doi.org/10.1038/nature17400

Zeeb, F. D., Baarendse, P. J. J., Vanderschuren, L. J. M. J., & Winstanley, C. A. (2015). Inactivation of the prelimbic or infralimbic cortex impairs decision-making in the rat gambling task. Psychopharmacology, 232(24), 4481–4491. https://doi.org/10.1007/s00213-015-4075-y

Zhao, Z., Han, X., Chen, R., Liu, Y., Bhattacherjee, A., Chen, W., & Zhang, Y. (2022). A molecularly defined D1 medium spiny neuron subtype negatively regulates cocaine addiction. Science Advances, 8(32), eabn3552. https://doi.org/10.1126/sciadv.abn3552

Zhukovsky, P., Puaud, M., Jupp, B., Sala-Bayo, J., Alsiö, J., Xia, J., Searle, L., Morris, Z., Sabir, A., Giuliano, C., Everitt, B. J., Belin, D., Robbins, T. W., & Dalley, J. W. (2019). Withdrawal from escalated cocaine self-administration impairs reversal learning by disrupting the effects of negative feedback on reward exploitation: A behavioral and computational analysis. Neuropsychopharmacology, 44(13), 2163–2173. https://doi.org/10.1038/s41386-019-0381-0

Zühlsdorff, K., Verdejo-Román, J., Clark, L., Albein-Urios, N., Soriano-Mas, C., Cardinal, R. N., Robbins, T. W., Dalley, J. W., Verdejo-García, A., & Kanen, J. W. (2024). Computational modelling of reinforcement learning and functional neuroimaging of probabilistic

reversal for dissociating compulsive behaviours in gambling and cocaine use disorders. BJPsych Open, 10(1), e8. https://doi.org/10.1192/bjo.2023.611