Title
Neural and behavioral features of flexible learning under uncertainty

Author
Aguirre, Claudia Gabriella

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Neural and behavioral features of flexible learning under uncertainty

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy in

Psychology

by

Claudia Gabriela Aguirre

2023

ABSTRACT OF THE DISSERTATION

Neural and behavioral features of flexible learning under uncertainty

by

Claudia Gabriela Aguirre

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2023

Professor Alicia Izquierdo Edler, Chair

We are surrounded by salient cues and actions in our everyday lives that predict reward in a constantly changing environment. A critical feature of goal-directed behavior is the ability to discriminate stimuli or actions that predict reward from those that do not, and further, to flexibly update our responses if predictions become inaccurate. This is referred to as behavioral flexibility, which measures our ability to re-evaluate previously learned associations that predicted reward and adjust our current responses following changes in the environment. Impairments in behavioral flexibility can lead to maladaptive behavior and impairments in decision-making, and is often associated with many neuropsychiatric disorders. The research presented here investigated the behavioral and neural basis of flexible learning under uncertainty by manipulating the associations and probability of reward outcomes in the environment, following alcohol experience and chemogenetic inhibition. Altogether, this approach establishes

a causal link between brain and behavior and increases our understanding of both adaptive and maladaptive decision-making.

We first tested the effect of prior alcohol exposure on flexible learning using a probabilistic, stimulus-based reversal learning task. Sex differences in alcohol consumption emerged, with females exhibiting heavier alcohol consumptions patterns than males. Furthermore, all alcohol-exposed animals, regardless of sex, were slower to learn, exhibited attentional deficits, and were less sensitive to negative feedback, compared to water-matched controls. Next, we compared the performance of rats on stimulus-based reversal learning in a rich (90/30) and poor (70/30) reward environment. All rats exhibited increased preservation during early reversal learning, resulting in slower learning. Additionally, a sex-dependent effect emerged on latencies to choose the better option, with females exhibiting longer latencies than males during reversal learning, an indication of slower processing speed. And finally, chemogenetic techniques were used to inhibit pyramidal neurons in the ventrolateral orbitofrontal cortex (vlOFC) and basolateral amygdala (BLA), on both a stimulus- and action-based reversal paradigm. Overall, the BLA seemed to be more necessary for learning action-based probabilistic reversals, as inhibition of this region resulted in poorer learning; whereas, the OFC was more necessary for detection of stimulus-based reversals. There was also evidence of sex-modulated learning flexibility, such that females with vlOFC inhibition were slower to learn during both deterministic and probabilistic action-based reversals compared to non-inhibited females; whereas, females with BLA inhibition were more impaired on probabilistic action-based reversals than non-inhibited females, this pattern was not observed in males. Altogether, these findings highlight the importance of including sex as a biological variable, as all of these studies found sex-dependent effects.

The dissertation of Claudia Gabriela Aguirre is approved.

Kate Wassum

Lara Allison Ray

Igor Spigelman

Alicia Izquierdo Edler, Committee Chair

University of California, Los Angeles

2023

Dedicated to my parents, Jose and Maria Aguirre, and sisters, Adriana, Blanca, Esther, and Isabel whose sacrifices, patience, unwavering support, and unconditional love, have been a constant source of inspiration and motivation throughout my academic journey.

Table of Contents

Table of Figures

Acknowledgements

Graduate school has been one of the most challenging, but rewarding experience of my life. I would not be where I am today without the support of my mentors, friends, and family. I would first like to acknowledge my advisor, Dr. Alicia Izquierdo, who beyond being a great scientist and role model, has also been an incredibly patient, thoughtful, supportive, encouraging, and generous mentor. Over the last six years Alicia has spent countless hours pushing me to my full potential, guiding me in experimental design, data analysis, and manuscript writing. I would also like to recognize my dissertation committee members: Dr. Kate Wassum, Dr. Lara Ray, and Dr. Igor Spigelman. Each one has challenged me to think critically about my project and provided invaluable advice throughout the years. A special thanks to my current and former lab mates, Juan Luis Romero Sosa, Valeria Gonzalez, Tony Ye, Zeena Rivera, and Alex Stolyarova for all of their help, encouragement, and support along the way, including all of my undergraduate research assistants, Kony, Johnny, Sriya, Lilly, and Madeline who were invaluable in helping me implement and complete all my experiments.

I would also like to thank my parents, Jose and Maria Aguirre, for all of their sacrifices, including immigrating to a new country, in order to ensure that my siblings and I would have better educational opportunities, which has allowed me to attain a higher education and pursue the career of my choice. To my older sisters, Blanca and Esther, who were the first in my family to go to college, for being good role models and motivating me to go to college. I am especially grateful for Blanca's encouragement to apply to research programs as an undergraduate student, as this would instill in me a love for research and shape the trajectory of my career. To Adriana, who helped raise me, and fostered my curiosity and love for reading and learning from an early age. To Esther, whose creativity and love for teaching, has shown me the importance of being a

good educator and using creative approaches to make learning more engaging at any age. To my little sister, Isabel, who is wise beyond her years, and has taught me to be more calm, patient, and empathetic.

Permissions

Chapter 2 is a version of Aguirre, C. G., Kanak, D., Stolyarova, A., Marty, V., Spigelman, I., Ray, L., & Izquierdo, A (2020). Sex-dependent effects of chronic intermittent voluntary alcohol consumption on attentional, but not motivational, measures during probabilistic learning and reversal. PLOS ONE 15(6): e0234729. https://doi.org/10.1371/journal.pone.0234729

Author contributions: C.G.A., Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; A.S., Data curation, Formal analysis, Software; K.D., Investigation; V.M., Data curation, Writing – review & editing; L.R. Conceptualization, Writing – review & editing; I.S, Funding acquisition, Resources, Writing – review & editing; A.I., Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing

Chapter 3 is a version of *Harris, C., *Aguirre, C.G., Kolli, S., Das, K., Izquierdo, A., & Soltani, A. (2021). Unique features of stimulus-based probabilistic reversal learning. Behavioral neuroscience, 135(4), 550–570. https://doi.org/10.1037/bne0000474 (*co-first authors)

Author contributions: C.G.A., Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing

– original draft, Writing – review & editing; C.H., Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing; S.K. and K.D., Investigation; A.I. and A.S., Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing **C.H. conducted all computational modeling for this study, including perseverative measures and parameters associated with reinforcement learning models (i.e. learning rate, inverse temperature). Refer to Appendix A for the results, methods, and figures generated by C.H.*

Chapter 4 was submitted to the Journal of Neuroscience and recently received a decision on major revision. We plan to resubmit after addressing reviewer comments. It is a version of this preprint:

Aguirre, C. G.*, Woo*, J.H., Sosa, J.L.R., Munier, J., Perez, J., Goldfarb, M., Das, K., Gomez, M., Ye, T., Pannu, J., Evans, K., O'Neill, P.R., Spigelman, I., Soltani, A., Izquierdo, A. (under review). Sex-dependent contributions of ventrolateral orbitofrontal cortex and basolateral amygdala to learning under uncertainty. https://www.biorxiv.org/content/10.1101/2023.04.03.

Author contributions: C.G.A., Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft; J.H.W, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft; J.L.R.S, Investigation, Software; T.Y., Software; J.P., K.D., M. Gomez, J.M., K.E., J.P., Investigation; M. Goldfarb, Data curation, Formal analysis, Software; P.R.O.N and I.S. Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization; A.I. and A.S., Conceptualization, Formal analysis, Funding

acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft.

**J.H.W. conducted all computational modeling associated with reinforcement learning models parameters (i.e. learning rate, inverse temperature, and decay parameter). Refer to Appendix C for the results, methods, and figures generated by J.H.W.*

Funding

The work conducted in Chapter 2 was supported by UCLA's Division of Life Sciences Recruitment and Retention fund (Izquierdo), R01 DA047870 (Izquierdo), and R01 AA024527 (Spigelman). The work conducted in Chapter 3 was supported by R01 DA047870 (Soltani and Izquierdo). We also acknowledge UCLA's Graduate Division Graduate Summer Research Mentorship and Graduate Research Mentorship programs (Aguirre) and the NSF Graduate Research Fellowship (Aguirre).

The work conducted in Chapter 4 was supported by UCLA's Division of Life Sciences Retention fund (Izquierdo), National Institutes of Health grants R01 DA047870 (Izquierdo and Soltani), R21 MH122800 (Izquierdo and Blair), R01AA024527 (Spigelman), K01 DA042219 (O'Neill), and F31 AA028183 (Munier), the NSF GRFP, Cota-Robles Fellowship, and Charles E. and Sue K. Young Fellowship (Aguirre), Ursula Mandel Fellowship (Aguirre, Romero-Sosa), and Graduate Research Mentorship award (Romero-Sosa), and the Training program in Neurotechnology Translation T32 NS115753 (Ye). We acknowledge the Staglin Center for Brain and Behavioral Health for additional support related to fluorescence microscopy. We also thank the NIDA Drug Supply program for the supply of clozapine-N-oxide.

Vita

Education

University of California, Los Angeles PhD Candidate, Psychology Ph.D. Program	2017-Present
University of Southern California Bachelor of Arts, Psychology	2010-2014

Awards and Honors

Charles E. & Sue K. Young Graduate Student Fellowship Award	2022-2023
SfN NIDA Diversity Scholar Travel Award Recipient	2018, 2022
National Hispanic Science Network (NHSN) Travel Award Recipient	2018-2022
Ursula Mandel Scholarship	2021-2022
UCLA BRI/Semel Institute's Graduate Travel Award	2021
Society for Neuroscience Scholars Program (NSP) Associate	2021
Atamdede Summer Research Fellowship	Summer 2020
Nancy Biram Summer Research Fellowship	Summer 2020
NSF Graduate Research Fellowships Program (GRFP) Recipient	2019-2022
UCLA Graduate Research Mentorship (GRM) Fellowship	2018-2019
UCLA Graduate Summer Research Mentorship (GSRM) Fellowship	Summer 2018, 2019
UCLA Cota-Robles Fellowship Recipient	2017-2018,2022-2023

Selected Publications

Woo, J.H., **Aguirre, C.G.,** Bari, B.A., Tsutsui, K.I., ,Grabenhorst, F., Cohen, J.Y., Schultz, W., Izquierdo, A., Soltani, A. (2023). *Mechanisms of adjustments to different types of uncertainty in the reward environment across mice and monkeys.* Cognitive, Affective, and Behavioral Neuroscience.

Nieto, S.J., Grodin, E.N., **Aguirre, C.G**., Izquierdo, A., Ray, L. *Translational opportunities in animal and human models to study alcohol use disorder*. Translational Psychiatry 11, 496 (2021).

*Harris, C., ***Aguirre, C.G**., Kolli, S., Das, K., Izquierdo, A., & Soltani, A. (2021). *Unique features of stimulus-based probabilistic reversal learning.*Behavioral neuroscience, 135(4), 550–570.(*co 1st authors)

Aguirre, C. G., Kanak, D., Stolyarova, A., Marty, V., Spigelman, I., Ray, L., & Izquierdo, A (2020). *Sex-dependent effects of chronic intermittent voluntary alcohol consumption on attentional, but not motivational, measures during probabilistic learning and reversal*. PLOS ONE 15(6): e0234729.

Selected Talks

Aguirre, C. G. (September 2022). *Roles of ventrolateral orbitofrontal cortex, basolateral amygdala and anterior cingulate cortex in flexible action-based learning.* Data Blitz presented at the 22nd Annual National Hispanic Science Network (NHSN) Scientific Meeting, Grand Rapids, MI.

Aguirre, C.G. (June 2021). California State University (CSU) Neuroscience Seminar: *Behavioral, computational, and neural basis of flexible stimulus-reward learning*, CSU, Long Beach.

Chapter 1: General Introduction

A critical feature of goal-directed, instrumental behavior is the ability to discriminate stimuli or actions that predict reward from those that do not, and further, to flexibly update the response to those stimuli/actions if predictions become inaccurate. We are surrounded by salient cues and actions in our everyday lives that predict reward in a constantly changing environment. For example, you might associate the famous golden arches with a juicy cheeseburger, or the green siren logo with a delicious cup of coffee. However, when your overconsumption of cheeseburgers and coffee results in high levels of cholesterol and gastrointestinal issues, we must override these previously rewarding associations that are now detrimental to your health, with alternate, healthier options, like sandwiches or smoothies. Behavioral flexibility refers our ability to re-evaluate these previously learned (stimulus or action-based) associations that predicted reward and adjust our current responses following changes in the environment in order to minimize risk and maximize reward (Izquierdo et al., 2017). Conversely, the inability to respond flexibly when stimuli or actions in our environment are no longer rewarding, can lead to maladaptive behavior and impairments in decision-making, also referred to as behavioral inflexibility or rigidity, and is a common feature across many neuropsychiatric disorders, including Alcohol and Substance Use Disorders (AUD; SUD), Attention Hyperactivity Disorder (ADHD), and Obsessive Compulsive Disorder (OCD) (Izquierdo & Jentsch, 2012; Uddin, 2021). Identifying the neural mechanisms underlying behavioral flexibility, requires the use of interference techniques that allow us to manipulate neuronal activity, on observable behavior associated with flexible learning and decision-making. This approach is ideal as it would establish a causal link between brain and behavior and increase our understanding of both adaptive and maladaptive decision-making.

Behavioral flexibility: Reversal Learning Paradigms

Traditionally, researchers have studied the neural substrates underlying behavioral flexibility across species using tasks collectively referred to as reversal learning paradigms. Importantly all of these paradigms are a good way to measure behavioral flexibility, but can vary greatly as there are many factors that can be manipulated in an environment. For example, you might want to choose a sensory modality that is representative to that species' natural environment (e.g., visual, olfactory, auditory), or vary the types associations, such as stimulus- or action-based associations, that they will need to learn and then remap following a reversal. Additionally, you can test for behavioral flexibility under varying levels of uncertainty, with tasks that involve deterministic or probabilistic reward outcomes, as well as manipulate the volatility in the environment, by decreasing or increasing the frequency that reversals occur. Reversal Learning paradigms have become the leading test for assessing behavioral flexibility for many important reasons: (1) high translatability across species, (2) easily adaptable to create varied environmental conditions, (3) often used to identify impaired cognitive processes in several neuropsychiatric disorders, including AUD and SUDs, OCD, Parkinson's, and schizophrenia (Brigman, Graybeal, et al., 2010; Finger et al., 2008; Izquierdo & Jentsch, 2012; Leeson et al., 2009; Remijnse et al., 2006; Swainson et al., 2000; van der Schaaf et al., 2013). Due to these factors and many more, it is not surprising that the use of these type of paradigms in published studies has increased exponentially in recent years (Izquierdo et al., 2017). Thus, it was imperative that I use reversal learning paradigms in my own studies in order to accurately assess behavioral flexibility under varying environmental conditions and following neural interference.

In all reversal learning paradigms, animals must first undergo discrimination learning, which in rodents, often involves pairing of a stimulus-based or spatial-based action (e.g., lever

pressing, digging in a bowl, nosepoking stimuli on a touchscreen, or displacing an object) with an outcome (e.g., a desirable food or water reward). In typical paradigms, two or more stimuli are presented concurrently and the subject learns about the features of the stimuli that bring about reward and those that do not (e.g., nosepoking S_A results in a better probability of reward than S_B; pressing left lever yields better payout than right lever; scent A is more rewarded than scent B) (Alvarez & Eichenbaum, 2002; Dalton et al., 2016; Eichenbaum et al., 1986; Izquierdo et al., 2013; Schoenbaum et al., 2000; Schoenbaum et al., 2002). With training, subjects become increasingly proficient at discrimination, in line with the associative rules imposed by the experimenter. The stimulus- or action-based outcome rules can be deterministic (e.g., S_A results in a sucrose pellet reward and S_B does not) or probabilistic (e.g., S_A results in a better probability of reward over S_B), with deterministic and probabilistic schedules of reinforcement producing marked dissociations in the neural circuitry recruited (Averbeck & Costa, 2017; Costa et al., 2016), and further in subsequent subsections.

In reversal learning paradigms (Izquierdo et al., 2017; Izquierdo & Jentsch, 2012), after either reaching a discrimination learning criterion for accuracy (Brushfield et al., 2008; Izquierdo & Jentsch, 2012; Stolyarova et al., 2019), a number of consecutive correct responses (Dalton et al., 2014; Dalton et al., 2016), or a fixed block length of trials (Farashahi et al., 2017; Soltani & Izquierdo, 2019a), the contingencies are reversed. At reversal, the trained response no longer results in a better probability of reward, though it usually remains the more frequently chosen option because of initial discrimination training. Indeed, usually reversals are acquired more slowly than the original discrimination (Brushfield et al., 2008; Schoenbaum et al., 2006). Perhaps partly due to this difference with original learning, reversal learning is considered unique in its requirement of flexibility, because it involves the subject inhibiting the prepotent response and,

instead, responding to stimuli that were previously irrelevant. We recently suggested that a thorough consideration of both accuracy and speed during discrimination and reversal phases can be particularly informative when studying neural signatures of flexible learning, as well as highlight potential impairments in attention, processing speed/deliberation, and motivation (Aguirre et al., 2020; Harris et al., 2021). Importantly, impairments in behavioral flexibility are commonly seen following chronic exposure to many drugs, including alcohol, and is often coupled with changes in neural activity in the fronto-amygdalar areas of the brain (Goudriaan et al., 2006; Koob & Volkow, 2010, 2016; Oscar-Berman et al., 2004).

Alcohol-Induced Changes in Behavioral flexibility

Alcohol Use Disorder (AUD) is considered a brain disease due to the neural disruptions and alterations caused by chronic use, but is also associated with behavioral inflexibility or rigidity as it is often characterized by persistent use despite negative consequences (Barker & Taylor, 2014; Trick et al., 2014; Wolff et al., 2018). This is of particular clinical relevance, given that inflexible decision-making, or the inability to flexible update responding following a contingency shift as measured by reversal learning and set-shifting tasks, has been shown to be a predictor of heightened self-administration of alcohol and cue-induced reinstatement of alcohol seeking (Laughlin et al., 2011; Loos et al., 2013), preference for alcohol, reduced latency to reach maximum alcohol consumption, and higher overall intake (Shnitko et al., 2019). Importantly, the effect of prior alcohol experience on behavioral flexibility has been more heavily studied than the bidirectional effect of behavioral flexibility on drinking outcomes (Melugin et al., 2021; Pitel et al., 2009; Stavro et al., 2013), and not many have extended this research by testing the possible neural mediators of either of these factors. Preclinical models can elucidate some of the underlying

mechanisms involved in cognitive deficits seen in individuals with AUD through the use of alcohol consumption models with high predictive validity, allowing for more precise examination of brain regions affected by alcohol use. Experimenters using rodents (mostly) have tested the effects of alcohol using paradigms assessing behavioral flexibility (Badanich et al., 2016; Fernandez & Savage, 2017; Fernandez et al., 2016), or one's ability to modify behavior based on changes in the environment.

Several groups have used preclinical models to test the effect of forced alcohol (i.e., ethanol, EtOH) exposure on reversal learning, a robust measure of behavioral flexibility (Izquierdo et al., 2017), that requires subjects to remap reward contingencies. Forced alcohol exposure models have tested the effects of alcohol on reversal learning using intragastric gavage (Badanich et al., 2016; Fernandez & Savage, 2017; Fernandez et al., 2016), intraperitoneal injections (Fernandez et al., 2016), and alcohol by passive vapor inhalation (Badanich et al., 2011). However, results have been mixed depending on method of administration, for example, intragastric gavages show reversal impairments in some studies (Badanich et al., 2016), but not others following abstinence (Fernandez et al., 2016), and have the added potential stressor of repeated oral intubation. However, vapor inhalation and intraperitoneal injections have both shown significant impairments in reversal learning following abstinence (Badanich et al., 2011; Fernandez et al., 2016). While exposure models ensure that rodents achieve higher blood alcohol concentrations and induce heavier alcohol dependence (Fidler et al., 2012; Karanian et al., 1986), their external validity is lacking compared to voluntary alcohol consumption models, which better model human consumption behavior (i.e., oral, voluntary consumption).

The most commonly used voluntary consumption model in rodents is the two-bottle choice (2-BC) procedure, which is a noninvasive self-administration method during which the animal is

given the choice to voluntarily consume ethanol and a non-ethanol beverage (usually water) orally, with either open or limited access (Leeman et al., 2010; Planeta, 2013) and was first developed and tested by Wise (1973) (Wise, 1973). The most common version of the two-bottle choice procedure is a chronic intermittent access schedule, in which rodents have 24-hour ethanol access, 3 days per week, followed by repeated periods of deprivation (usually every other day) (Charlton et al., 2019; Fisher et al., 2017; Spoelder et al., 2017). This rapid alternation between drinking and non-drinking days has been shown to reliably increase ethanol intake over time in both rats (Hargreaves et al., 2009; Loi et al., 2010; Rodd-Henricks et al., 2000; Simms et al., 2008; Wise, 1973) and mice (Hwa et al., 2011; Melendez, 2011).

Although there have been several studies using voluntary alcohol consumption models to assess for potentially deficits in several types of decision-making involving risk-preference (McMurray et al., 2016; Miller et al., 2017; Spoelder et al., 2017), delayed discounting (Kruse et al., 2017), and delayed reward (Spoelder et al., 2017), only a few groups have tested the effect of alcohol on reversal learning in rodents (Aguirre et al., 2020; Charlton et al., 2019; Fisher et al., 2017). Even though all of these studies used the same alcohol concentration (20%) and a chronic intermittent schedule (i.e., 3-24 hr. days of alcohol access), the findings have been somewhat mixed, perhaps due to differences in duration of alcohol administration period, sex, and deterministic vs. probabilistic nature of the reversal task. Fisher (2017) found no group (alcohol vs. control) differences in performance on reversal learning, which may have been attributed to a shorter duration of access and a sharp de-escalation instead of escalation of alcohol consumption over time. Conversely, Charlton et al., (2019) found deficits in reversal learning (i.e., fewer correct trials, higher perseverative index, more session to reach criterion than controls) after 26 weeks of alcohol access, but only during early learning (i.e., before mid-criterion of 50% correct was

reached), after which these deficits were overcome. I also found deficits (i.e., more initiation omissions, slower initiation latencies) during early learning (i.e., pre-training and early discrimination learning), but they did not persist into reversal learning (Aguirre et al., 2020). However, it is important to note that I used a probabilistic reversal learning task (i.e. 70/30) in my original study, unlike the aforementioned studies, which made it increasingly more difficult for all groups to learn even after extensive testing (both alcohol and controls), and used both male and female rats, which yielded a sex-dependent difference in alcohol consumption, with females escalating their consumption at a faster rate than males, consistent with prior literature (de la Torre et al., 2015; Hwa et al., 2011; Vetter-O'Hagen et al., 2009; Vetter-O'Hagen & Spear, 2011).

Sex differences in alcohol consumption patterns

Several preclinical studies that included females have found they seem to be more vulnerable to the effects of alcohol, resulting in greater overall consumption and escalation over time (de la Torre et al., 2015; Hwa et al., 2011; Vetter-O'Hagen et al., 2009; Vetter-O'Hagen & Spear, 2011). Some groups have suggested that increases in alcohol intake over time may be hormone-dependent (Torres et al., 2014), which may be contributing to the enhanced rewarding effects of alcohol in females compared to males. However, a systematic review conducted on the effects of hormones on alcohol consumption using both human and animal literature yielded inconclusive results, as some studies found an association between hormone levels and alcohol intake, while others did not (Erol et al., 2019).

Orbitofrontal Cortex (OFC) function in flexible learning

Several brain regions have been identified as being involved with different aspects and types of flexible learning and decision making, including the posterior ventral orbitofrontal cortex (OFC). The OFC is divided into six main subregions (i.e., ventrolateral orbital (VLO), ventral orbital (VO), lateral orbital (LO), medial orbital (MO), dorsolateral orbital (DLO), agranular insular (AI) areas); however, for the purpose of our experiments, we will be focusing on the ventrolateral subregions of the OFC (i.e., VLO, VO, LO). The overlap in OFC-striatal organization between rodents and non-human primates has long been established, specifically the OFC's connection to the mediodorsal thalamus (Groenewegen, 1988; Ongür & Price, 2000; Ray & Price, 1992; Uylings & van Eden, 1990); however, our primary focus will be on its connections to other frontocortical regions, like the basolateral amygdala. Studies have shown that the MO and VO subregions of OFC both project to the anterior cingulate, while the MO sends projectors to a wider range of limbic areas, like the basolateral and central amygdala, than does the VO (Hoover & Vertes, 2011). However, the lateral subregions of the OFC seem to be more involved with sensory integration, like the formation and maintenance of stimuli to outcomes, which is of particular interest in my experiments.

Stimulus-Based Learning. Substantial evidence indicates the ventral OFC is engaged when stimulus-based contingencies become stable with experience (Riceberg & Shapiro, 2012, 2017). The OFC is believed to play a role in forming stable stimulus-based representations by encoding reward predictions, supporting learning of the values of options (Steiner & Redish, 2012; Sul et al., 2010), comparing values before and after making a choice (Padoa-Schioppa & Assad, 2006; Padoa-Schioppa & Conen, 2017), and updating values, as necessary (Baltz et al, 2018). Animals with lesions in, or pharmacological inactivations of, OFC exhibit impairments in using consistent strategies (Robbins & Cardinal, 2019; Verharen et al., 2020), showing a tendency to

switch more frequently between choices, demonstrating a difficulty in assigning reward value to choices (Noonan et al., 2012). However, other studies have shown that OFC inactivation increases both perseverative (i.e. inability to shift choice) and regressive errors (i.e. shifting choice even if previously reinforced), increasing the number of trials required to reach criterion in reversal learning (Alsiö et al., 2015; Brigman et al., 2013; Chudasama & Robbins, 2003; Dalton et al., 2016; Kim & Ragozzino, 2005; Ragozzino, 2007), suggesting impairments in the ability to maintain a new choice and flexibly shift to new strategy (Churchwell et al., 2009). Furthermore, vOFC neurons respond to the expected outcome and track reward value across reversal learning (Moorman & Aston-Jones, 2014), with prior studies showing OFC impairs reversals based on visual stimuli (Izquierdo et al., 2013), and spatially-cued operant responses (Boulougouris et al., 2007; Mar, 2011). However, there is more emphasis on cue-guided learning (i.e., stimulus-based) in VO and VLO studies, with few investigations probing their role in action-outcome coding.

Action-Based Learning Although the OFC has mostly been implicated in detecting changes in stimulus-reward contingencies (in nonhuman primates, mostly), its role in the encoding of actions associated with rewards is less clear and has yielded mixed results. Several studies have found that both OFC-lesioned and OFC-inactivated animals, are unable to appropriately update their actions (e.g. reduce lever pressing) once a particular action's associated reward has been devalued, suggesting that the OFC conveys information about action-value and is involved in detecting changes in reward value and adjusting instrumental actions accordingly (Balleine et al., 2011; Gremel & Costa, 2013; Ostlund & Balleine, 2007b; Parkes et al., 2017), with some only observing this if pre-training was included (Parkes et al., 2017). Conversely, studies using the Pavlovian-Instrumental-Transfer (PIT) task, which requires the animal to form both S-O and A-O associations and use that knowledge to guide behavior, often following devaluation, have found

that OFC lesions made after training abolished S-O transfer, but did not affect R-O devaluation (Ostlund & Balleine, 2007a; Panayi & Killcross, 2018; Pickens et al., 2005). Spatial-based reversal learning paradigms, which require the animal to form and flexibly update action-outcome associations, have yielded similar findings in that OFC-lesioned animals perform poorly following a reversal (i.e., made more incorrect responses and perseverative errors) initially, but not after the new action-reward associations have been formed (Boulougouris et al., 2007). Interestingly, these same impairments were not found if the lesions were done after pre-surgical training (Boulougouris & Robbins, 2009), suggesting that the OFC is not needed after associations has been formed with prior task experience, but is necessary when task demands are high, such as following a reversal. It is important to note that results may differ depending on whether a reversal learning task is deterministic or probabilistic, given that some studies have found that inactivation of lOFC increases the number of trial omissions and response latencies (Dalton et al., 2016), and decreases the number of reversals (Verharen et al., 2020) in a probabilistic reversal learning task, whereas none of these deficits were observed following reversals in a deterministic task with assured outcomes (Dalton et al., 2016).

Overall, these findings suggest that the OFC plays an important role in detecting contingency changes in either stimulus- or action-outcome associations after either devaluation or reversals, but is still predominantly involved in encoding and updating stimulus-based associations contingencies.

Basolateral amygdala (BLA) function in flexible learning

The Basolateral Amygdala (BLA) is comprised of glutamatergic pyramidal neurons (~85%), which are the primary projection neurons (Klenowski et al., 2015; McDonald, 1982, 1992), these

neurons receive glutamatergic inputs from cortical and subcortical regions, carrying a variety of information (e.g. sensory , executive, and memory) (Cassell & Wright, 1986; Ottersen, 1982; van Vulpen & Verwer, 1989), and then transmit this information downstream to other brain regions to guide associative memories, particularly in response to emotionally relevant stimuli (Daviu et al., 2019; Lalumiere, 2014; Phelps & LeDoux, 2005). The remaining neuronal population in the BLA is made up of GABAergic interneurons, which modulates surrounding pyramidal neurons (Carlsen, 1988; Klenowski et al., 2015; Spampanato et al., 2011).

Stimulus-Based Learning. Although the BLA has traditionally and extensively been studied in the context of fear conditioning and fear learning (Fanselow & Gale, 2003; Maren & Fanselow, 1996), it seems to not only be required in integrating the sensory properties of aversive stimuli with their affective valence, but also appetitive stimuli (Fanselow & LeDoux, 1999; Ghashghaei & Barbas, 2002; Lüthi & Lüscher, 2014; Morrison & Salzman, 2010). In fact, BLA neurons have been shown to respond to both cues and rewards during appetitive Pavlovian conditioning, and thus are critical for the encoding of these stimulus-based associations, used to guide subsequent action selection (Malvaez et al., 2015; Sias et al., 2021), but also the retrieval of these associations (Malvaez et al., 2019); these findings are consistent with prior evidence showing that BLA lesions and inactivation disrupt stimulus-based associations in rats (Corbit & Balleine, 2005; Derman et al., 2020; Lichtenberg et al., 2017; Malvaez et al., 2015; Ostlund & Balleine, 2008). The BLA also seems to be involved in encoding action-outcome contingencies and mediates the influence of outcome value on future instrumental actions (Balleine et al., 2003; Corbit & Balleine, 2005; Wang et al., 2005), thus suggesting the BLA might be playing a more general role in outcome encoding that is non-specific to either stimulus or action- outcome associations.

Although the role of the BLA in behavioral flexibility has not been as extensively studied, nevertheless, there is evidence to suggests it might also be involved given that it is required for responding to violations in a cue or action's recent reward history (e.g., contingency shifts) to the extent that the response relies on an outcome-specific representation for the comparison (Paton et al., 2006; Schoenbaum et al., 1999). Although it is generally not necessary for the formation of initial stimulus-reward associations during discrimination learning (Schoenbaum et al., 2003), selective pre-training BLA lesions have been shown to slightly enhance the ability to form stimulus-reward associations following a reversal (Izquierdo et al., 2013), with inactivation studies showing this enhancement only during late reversal learning (Hervig et al., 2019). However, the BLA's role in reversing these associations is even less clear, with some studies showing impairments (Churchwell et al., 2009; Schoenbaum et al., 2003), no effect (Murray & Izquierdo, 2007), and a nullification of learning deficits produced by OFC lesions after BLA lesions (Stalnaker et al., 2007). Furthermore, selective BLA lesions facilitate responding after negative feedback in rats (Izquierdo et al., 2013).

Overall, these findings suggest that the BLA is heavily involved in updated outcome-specific representations after there has been a change in reward value, which seems to be non-specific to either stimulus-based or action-based associations.

Sex differences in strategies

Orsini et al. (2017) did an extensive review on sex differences in animal models of decision-making, and found that the strategies that males and females use depend on the type of uncertainty and risk modeled in the environment. In tasks in which the probability distribution is unknown, female rodents tend to switch more between options before deciding which is the most

optimal choice, while male rodents seem to use more global information when making a decision. Conversely, in tasks with a known probability distribution, females develop an optimal choice more quickly than males (Orsini et al, 2017). Additionally, recent papers from the Grissom lab have found that although males and females achieved the same level of performance in terms of accuracy, there are sex-dependent strategies that emerged (Chen et al., 2021a, 2021b), such that female mice displayed more win-stay behaviors than males, as they were more likely to choose the same option that was rewarded on a previous trial, learn quicker during exploration and exploit the better option much quicker than males.

Reinforcement Learning Models

Reinforcement Learning (RL) models are commonly used to model adaptive decision-making and behavior in response to rewards and punishment. They allow us to better understand how an animal and/or person can optimize their behavior in order to maximize their rewards in an environment, either natural or artificial, with varying degrees of uncertainty (expected or unexpected) and volatility (Dayan & Niv, 2008; Soltani & Izquierdo, 2019; Soltani & Koechlin, 2021). Specifically, they provide a framework through which we can understand how different learning strategies are used in order to associate stimuli, actions, and outcomes, and guide our future choices and behavior. However, reinforcement learning models provide more than just a computational framework for stimulating and predicting learning and decision-making, they are also used to stimulate and make predictions about underlying neural mechanisms that drive this behavior. Evolutionarily, mammalian brains have evolved in a way that allows us to flexibly adapt to different and changing environments, which is thought to require the PFC (Soltani & Koechlin, 2021). As discussed previously, many regions of the PFC in both primates and rodents have been

identified as being crucial for adaptive learning and decision-making, including the OFC and ACC. Thus, it is of increasing importance to use computation models, like reinforcement learning models, to better understand how these computations are made in the brain.

Currently, there are a variety of RL models used that mostly fall into two categories (model-free and model based) and have been used to simulate and predict reward learning behavior (Sutton & Barto, 1998). A model-based RL framework uses representations and expectations of the environments to predict the consequences of one's actions (e.g. rewards or states), such that new information can be evaluated by simulating behavioral trajectories based on prior experience, instead of through trial and error; whereas, a model-free RL framework uses prediction errors to progressively acquire cache estimates of the value of prior choices and actions through trial and error (Dayan & Berridge, 2014; Dayan & Niv, 2008; Sutton & Barto, 1998). Thus, although model-based RL models allow for greater behavioral flexibility than model-free RL models, this is at the expense of greater computational complexity. However, even these traditional frameworks have their limitations and fail to capture certain behavior, thus new models, like successor representation that integrate both of these models (Gershman, 2018; Lehnert & Littman, 2019), along with new entropy-based metrics (Trepka et al., 2021; Xin et al., 2020), can be used to more accurately predict behavior.

Our lab, in collaboration with the Soltani lab, has found that a reinforcement model containing a single learning rate parameter and inverse temperature parameter, better captured stimulus-based discrimination and reversal learning in animals (Harris, Aguirre et al., 2021). We found that contingency changes following a reversal elicited higher learning rate and reduced sensitivity to differences in reward value (i.e., inverse temperature), compared to discrimination learning, suggesting faster learning and more explorative behavior following a reversal. Recent

work has also confirmed that this model is better at capturing stimulus-based learning, than action-based learning after a reversal, such that a higher learning rate and reduced sensitivity to differences in reward value was found during the stimulus-based task relative to the action-based task, suggesting action learning elicits more exploitative choice behavior and less of a need to frequently update reward value (Aguirre et al., unpublished).

Chapter 2: Sex-dependent effects of chronic intermittent voluntary alcohol consumption on attentional, not motivational, measures during probabilistic learning and reversal

ABSTRACT

Forced alcohol (ethanol, EtOH) exposure has been shown to cause significant impairments on reversal learning, a widely-used assay of behavioral flexibility, specifically on fully-predictive, deterministic versions of this task. However, previous studies have not adequately considered voluntary EtOH consumption and sex effects on probabilistic reversal learning. The present study aimed to fill this gap in the literature. Male and female Long-Evans rats underwent either 10 weeks of voluntary intermittent 20% EtOH access or water only (H2O) access. Rats were then pretrained to initiate trials and learn stimulus-reward associations via touchscreen response, and subsequently required to select between two visual stimuli, rewarded with probability 0.70 or 0.30. In the final phase, reinforcement contingencies were reversed. We found significant sex differences on several EtOH-drinking variables, with females reaching a higher maximum EtOH consumption, exhibiting more high-drinking days, and escalating their EtOH at a quicker rate compared to males. During early abstinence, EtOH drinkers (and particularly EtOH-drinking females) made more initiation omissions and were slower to initiate trials than H2O drinking controls, especially during pretraining. A similar pattern in trial initiations was also observed in discrimination, but not in reversal learning. EtOH drinking rats were unaffected in their reward collection and stimulus response times, indicating intact motivation and motor responding. Although there were sex differences in discrimination and reversal phases, performance improved over time. We also observed sex-independent drinking group differences in win-stay and lose-shift strategies specific to the reversal phase. Females exhibit increased vulnerability to EtOH effects in early learning:

there were sex-dependent EtOH effects on attentional measures during pretraining and discrimination phases. We also found sex-independent EtOH effects on exploration strategies during reversal. Future studies should aim to uncover the neural mechanisms for changes in attention and exploration in both acute and prolonged EtOH withdrawal.

INTRODUCTION

Individuals with Alcohol Use Disorder (AUD) show cognitive impairments, particularly in the domain of behavioral flexibility, broadly defined as the ability to adjust one's behavior in response to changes in the environment (Dajani et al., 2015). Preclinical models that mimic alcohol (ethanol, EtOH) consumption in humans can elucidate underlying mechanisms related to such cognitive impairments in individuals with AUD (Houston et all, 2014).

Several groups have tested the effect of forced EtOH exposure on reversal learning, a robust measure of behavioral flexibility commonly used in experimental animals (Izquierdo et al., 2017) that requires the remapping of reward contingencies. In deterministic (fully-predictive) reversal learning paradigms more frequently employed in behavioral pharmacology experiments, subjects first learn to discriminate and choose between two stimuli, one of which is rewarded and the other which is not. After successful discrimination, the associated outcomes of the two stimuli are reversed, forcing the subject to remap the associations (Izquierdo et al., 2017). In probabilistic reversal learning (PRL) paradigms, each stimulus is associated with a probability of reward (e.g. .80/.20, .70/.30), with one stimulus associated with a higher probability of reward (the "better" option), and another with a lower probability of reward (the "worse" option) (Izquierdo et al., 2017; Amodeo et al., 2017; Dalton et al., 2016).

Voluntary alcohol consumption models (e.g., 2-Bottle Choice, EtOH gelatin) have been used to assess effects on deterministic reversal learning (Fisher et al., 2017; Ray et al., 2018; McMurray et al., 2014), resulting in no significant treatment group differences. However, it is worth noting that these studies only examined overall performance (i.e., trials to reach criterion, number of correct choices and errors) and did not report whether there were group differences on latency measures to initiate and commit trials and collect reward, or the types of strategies

employed on a more fine-grained, trial-by-trial basis. Latency measures, for example, could be used to dissociate processing speed and decision-making from motivational effects. Latency to collect reward along with number of initiated trials are commonly used as a measure of motivation, whereas response latency is often used as a measure of processing or decision-making speed in the Five-Choice Serial Reaction Time Task (5CSRTT) (Amitai et al., 2010; Asinof et al., 2014; Bushnell et al., 2009; Remmelink et al., 2017; Guidi et al., 2015; Marbach et al., 2017). Measures of attention vary, but can include number of omissions (i.e., failure to make a response before the end of the trial), as well as percentage of correct responses (i.e., accuracy) (Amitai et al., 2010; Bushnell et al., 2009; Remmelink et al., 2017; Bruinsma et al., 2019; Martin et al., 2015; Bayless et al., 2012). Although omissions are thought to be indicative of inattentiveness, one must consider this measure in conjunction with stimulus response and reward collection latencies to rule out processing speed and motivational deficits, respectively, as potential confounds (Bayless et al., 2012; Turner et al., 2016). In particular, a failure to initiate trials, or taking longer to do so, could be indicative of a deficit in task engagement, and when observed in parallel with intact reward collection or stimulus-response times, could rule out deficits in motivation to learn about stimuli that predict rewards and to procure rewards.

We also studied Win-Stay/Lose-Shift (WSLS) strategies commonly analyzed in PRL paradigms. WSLS strategies reflect an animal's tendency to select the same stimulus after being rewarded (i.e., Win-Stay) or switch to select a different stimulus after not being rewarded (i.e. Lose-shift) (Worthy et al., 2014). Conversely, animals may use less advantageous strategies, such as selecting a different stimulus after being rewarded (i.e., Win-Shift), or choosing the same stimulus after not being rewarded (i.e. Lose-Stay). We probed these in the present study. Also noteworthy, most studies have exclusively used male animals, limiting the generalizability of the

results given recent findings showing sex differences in consumption patterns, with female rodents showing higher EtOH intake levels and preference for EtOH (Vetter-O'Hagen et al., 2009, 2011; Lourdes de la Torre et al., 2015; Hwa et al., 2011; Wallin-Miller et al., 2017), even exhibiting less aversion to EtOH compared to males (Cailhol et al., 2002; Sherrill et al., 2011; Schramm-Sapyta et al., 2014; Marquardt et al., 2017).

The present study sought to address these gaps in the literature by probing the effects of a chronic intermittent voluntary alcohol consumption model on PRL in male and female rats. Rats were administered a 2-bottle choice procedure, during which they were given access to either both 20% EtOH and H2O, or H2O only, 3 days per week for a total of 10 weeks. Five days after their last day of EtOH access they underwent pretraining and advanced to PRL after meeting several training criteria. Given recent findings on sex differences in EtOH consumption, we hypothesized that females would be more EtOH-preferring and reach higher EtOH consumption levels than males. Here, we corroborate previous findings of enhanced EtOH consumption and escalation in female rats compared to male rats. Surprisingly, we found sex-dependent treatment group differences in trial initiations, with the most robust effects in pretraining that carried through early discrimination. These results were in contrast to intact committed trials (i.e., trials in which animals responded to the presented stimuli) and reward collection times throughout learning. We found changes in WSLS strategy specific to the reversal phase, with EtOH drinkers displaying more exploration (i.e. shift) strategies than H2O drinkers. These effects were not sex-dependent. Taken together, the present results suggest an enduring effect on attention and exploration-based strategies, but not motivational measures, in stimulus-reward association learning in prolonged abstinence from EtOH.

RESULTS

EtOH consumption

Independent samples t-tests showed that females reached a greater maximum level of EtOH consumption [t(14)=3.46, p=0.004] (**Figure 2-2A**) and exhibited more high-drinking days (i.e. days of EtOH consumption 5+ g/kg/24 hours) [t(14)=3.00, p=0.01] than males (**Figure 2-2B**). A repeated-measures ANOVA was used to assess the within-subject effect of EtOH drinking days and the between-subject factor of sex on EtOH consumption (g/kg). There was a within-subject effect of day [F(28, 392)=6.68, p<0.0001], suggesting an escalation of EtOH over the course of 29 EtOH drinking days (**Figure 2-2C**). A marginally-significant sex*day interaction was found, [F(28,392)= 1.51, p=0.05], with female animals escalating drinking more steeply than males over the 29 days. Although not significant, we found a trend for a main effect of sex [F(1, 14)=3.42, p=0.09]. Conversely, we saw a de-escalation of water (H2O) consumption over the 29 days of alcohol access, [F(28,392)=13.594 p<0.0001; **Figure 2-2D**], but found no sex*day interaction, [F(28,392)= 0.993, p=0.48], yet a trend for an effect of sex [F(1, 14)=3.25, p=0.09]. It is important to note that H2O-drinking animals reached a significantly higher body weight 413.44±20.87 (M±SEM), compared to EtOH-drinking animals 332.69±24.09, [t(30)=2.53, p=0.02], and that body weight was significantly negatively correlated with maximum EtOH consumption level [r(16)=-0.67, p=0.005], and number of high-drinking days [r(16)=-0.61, p=0.01]. Overall females exhibited heavier EtOH consumption patterns than males.

Learning data were analyzed separately, subdivided into pretraining, discrimination, and reversal phases. Early vs. late discrimination and reversal learning based on prior studies indicating these may be particularly informative to contrast (Piantadosi et al., 2019; Bryce et al., 2015; Stolyarova et al., 2019; Izquierdo et al., 2010, 2017; Robbins et al., 2019).

Pretraining (PT) performance

As pretraining was a unique phase with slightly different dependent measures (i.e. forced choice omissions and latencies), a separate GLM was used to assess the effect of group (EtOH, H2O) and sex (female, male), and group*sex interactions on the number of pretraining sessions required to reach criterion to advance to the discrimination learning phase of the PRL task. A significant effect of group and sex emerged, with the EtOH-experienced animals requiring a greater number of pretraining sessions than the H2O-only group (GLM: βgroup =-14.00, p=0.04; **Figure 2-3A**), and females requiring a greater number of pretraining sessions than males (GLM: βsex =-11.38, p=0.001; **Figure 2-3A**). A group*sex interaction was found (GLM: βgroup*sex =12.00, p=0.01) with EtOH-drinking females requiring more pretraining sessions than EtOH-drinking males (GLM: βsex =-11.38, p=0.03). Both EtOH-drinking females (GLM: βgroup =-14.00, p=0.01) and EtOH-drinking males (GLM: βgroup =-2.00, p=0.03) needed more sessions to reach criterion than their H2O-drinking counterparts (**Figure 2-3A**). There was an overall average of 10.19±1.47 (M±SEM) sessions to successfully meet the pretraining criterion and advance to discrimination learning.

We analyzed group and sex differences, as well as the group by sex interaction on initiation omissions, defined as a failure to nosepoke the center white square stimulus within 40s, and initiation latencies, defined as the duration until nosepoke of the center white square stimulus. There was an effect of both group (GLM: βgroup = -554.12, p=0.0001; **Figure 2-3B**), and sex (GLM: βsex = -517.87, p=0.0003; **Figure 2-3B**) on initiation omissions, with EtOH-drinking animals and females displaying more initiation omissions compared to H2O-drinking animals and males, respectively. A significant drinking group*sex interaction (GLM: βgroup*sex = 467.00, p=0.01; **Figure 2-3B**), indicated that both EtOH-drinking females (GLM: βgroup =-554.00

p=0.01) and males (GLM: βgroup = -87.13, p=0.01) exhibited more initiation omissions than their H2O-drinking counterparts, and EtOH-drinking females exhibited more initiation omissions than EtOH-drinking males (GLM: βsex = -517.88, p=0.01). The EtOH group also displayed longer initiation latencies than the H2O group (GLM: βgroup = -3.48, p=0.001; **Figures 2-3C, 2-10A**), but there was no effect of sex (GLM: βsex = -1.37, p=0.18; **Figures 2-3C, 2-10D**), or group*sex interaction (GLM: βgroup* sex = -0.09, p=0.95; **Figure 2-3C**).

Next we analyzed differences in forced-choice omissions, defined as failure to nosepoke the stimulus presented on either the left or right side of the touchscreen after initiation of the trial, and forced-choice latencies, defined as duration until nosepoke of the stimulus presented on the left or right side. There was a significant effect of sex (GLM: βsex = -6.13, p=0.03; **Figure 2-3D**), with females displaying more forced-choice omissions than males, but no significant group (GLM: βgroup = -4.25, p=0.13; **Figure 2-3D**), or group*sex interaction (GLM: βgroup*sex = 5.38, p=0.17; **Figure 2-3D**). There was a significant effect of both group (GLM: βgroup = -2.76, p=0.01; **Figures 2-3E, 2-10B**), and sex (GLM: βsex =-2.37, p=0.03; **Figures 2-3E, 2-10E**), on forced-choice latency, with the EtOH-experienced animals and females taking longer to select the stimulus compared to H2O-drinking animals and males, but no significant group*sex interaction (GLM: βgroup*sex = -0.45, p=0.76). Finally, we analyzed data for differences in reward latency, defined as the duration to collect the sucrose reward from the food magazine, and found no effect of group (GLM: βgroup = 0.44, p=0.70; **Figure 2-3F, 2-10C**), sex (GLM: βsex =-0.01, p=0.92; **Figures 3F, 2-10F**), or group*sex (GLM: βgroup*sex = 0.15, p=0.35; **Figure 2-3F**) interaction.

Collectively, the omission and latency data for initiations and forced-choice trials suggest an attenuating effect of EtOH experience, and specifically in EtOH- experienced females, on

quickly responding to stimuli, while the reward collection data point to preserved motor responding and motivation for reward in EtOH-experienced animals.

Probabilistic Discrimination (D) performance

There were no significant group (GLM: βgroup =-9.50, p=0.07) or sex differences (GLM: βsex = 3.13, p=0.54) on total number of sessions to reach criterion for the discrimination phase, but a marginally-significant group*sex interaction on this measure was observed (GLM: βgroup*sex =14.88, p=0.047; **Figure 2-4A**). Overall, all animals performed comparably regardless of group or sex, with an overall average of 27.41±2.17 days required to successfully discriminate and advance to the reversal phase.

A GLM model was used to test the effect of drinking group (EtOH, H2O), sex (female, male), days, and their 2-way and 3-way interactions on probability correct (i.e. choosing the better option), number of rewards (i.e. sucrose pellets), and initiation omissions across 20 testing days of discrimination learning (D1-D20). All animals demonstrated learning by showing an increase in choosing the better option across days, (GLM: βday =0.01, p=0.001; **Figure 2-4B**), regardless of group or sex. There was no effect of drinking group (GLM: βgroup = -0.01, p=0.79), or sex (GLM: βsex =0.07, p=0.17), and no group*sex (GLM: βgroup* sex = -0.03, p=0.65), or sex*day (GLM: βsex* day =0.01, p=1.00) interaction on probability correct. There was however a significant group*day interaction (GLM: βgroup*day = 0.02, p=0.02), with the H2O-drinking animals choosing the better option increasingly more across days than the EtOH-drinking animals. There was also a significant group*sex*day interaction (GLM: βgroup*sex*day = -0.02, p=0.01). Post-hoc comparisons revealed males learned quicker than females (GLM: βsex= 0.01, p<0.001), that H2O-drinking females chose the better option increasingly more across days than their EtOH-

drinking counterparts (GLM: βgroup =0.01 p=0.002); and that the same pattern was not observed for males.

All animals increased the number of rewards collected across days, (GLM: βday =1.83, p=<0.0001; **Figure 2-8A**). Drinking group (GLM: βgroup=22.59, p=0.01), and sex differences (GLM: βsex =17.60, p=0.04) also emerged, with both H2O-drinking animals and males collecting a greater number of rewards than EtOH-drinking animals and females, respectively (**Figure 2-8A**). No significant group*sex (p=0.07), group*day (p=0.78), or group*day*sex (p=0.08) interactions on number of rewards collected were found, only a significant sex*day interaction (GLM: βsex*day =-1.75, p<0.0001), with females displaying a greater increase in rewards collected across days (GLM: βday =1.69, p<0.0001), compared to males (**Figure 2-8A**).

Conversely, all animals decreased the number of initiation omissions across days, (GLM: βday =-0.29, p=0.02; **Figure 2-8B**), regardless of group or sex. There were significant group (GLM: βgroup =-11.17, p=0.001) and sex differences (GLM: βsex =-6.36, p=0.04), with EtOH-drinking animals and females exhibiting more initiation omissions than H2O-drinking and males, respectively (**Figure 2-8B**). A significant group*sex interaction (GLM: βgroup*sex =14.81, p=0.01), with EtOH-drinking females displaying more initiation omissions than H2O-drinking females (GLM: βgroup =-304.13, p=0.02); the same pattern was not observed in males (**Figure 2-8B**). Similarly, a significant group*sex*day (GLM: βgroup*sex*day =-0.69, p=0.02) revealed that EtOH-drinking females decreased their number of initiation omissions across discrimination learning at a slower rate than EtOH-drinking males (p=0.03) and H2O-drinking females (p=0.03). There were no differences between male and female H2O drinkers, or between EtOH- vs. H2O-drinking males.

Finally, we conducted GLM analyses on the sum of initiation and choice omissions and median (initiation, correct/incorrect choice, reward) latencies in discrimination learning, collapsed across days of testing. Our results indicated EtOH-drinking animals exhibited more initiation omissions than H2O-drinking animals (GLM: βgroup =-304.12, p=0.01), whereas H2O-drinking animals exhibited longer reward collection latencies than EtOH-drinking animals (GLM: βgroup=0.17, p=0.02). All other types of omission and latency analyses yielded non-significant results.

First 500 trials. Because interesting latency differences were obtained for the pretraining phase and based on prior studies comparing early vs. late discrimination learning, we conducted further latency analyses for other phases of learning to assess if these trends were maintained. An analysis of initiation latencies and omissions, choice (correct and incorrect) latencies and omissions, and reward latencies was conducted for the first 500 trials of discrimination learning to capture early learning in this phase, with animals averaging 89.19±3.38 committed trials per day. There was a significant group difference on initiation omissions (GLM: βgroup =-158.75, p=0.002; **Figure 2-4C**), with EtOH-drinking animals exhibiting more initiation omissions than H2O-drinking animals, but no sex differences (GLM: βsex = -56.50, p=0.36; **Figure 2-4C**). A significant drinking group*sex interaction on initiation omissions emerged (GLM: βgroup*sex = 219.00, p=0.02; **Figure 2-4C**), with EtOH-drinking females exhibiting more initiation omissions than their H2O-drinking female counterparts (p=0.01), and H2O-drinking males exhibiting marginally more initiation omissions than H2O-drinking females (p=0.05), but EtOH-drinking males were no different than H2O-drinking males. There was no effect of drinking group (**Figures 2-4D, 2-11A**), sex (**Figures 2-4D, 2-11E**), or group*sex interaction (**Figure 2-4D**) on initiation latencies, choice omissions, correct or incorrect choice latencies, or reward collection

latencies (**Figures 2-11B, 2-11F**), with the exception of a marginally significant group effect (GLM: βgroup =0.88, p=0.05), with H2O-drinking animals exhibiting more choice omissions than EtOH-drinking animals. It should be noted that choice omissions represented a small number of occurrences (normally ranging from 0-2) and this effect was driven by a single outlier (>2 SD from the mean) with 4 choice omissions, which upon removal yielded a non-significant group effect (GLM: βgroup =0.48, p=0.14). In summary, the data for early discrimination learning suggest an enduring effect of EtOH experience on initiating trials, with females most affected.

Last 500 trials. An analysis of initiation latencies and omissions, choice (correct and incorrect) latencies and omissions, and reward latencies was also conducted for the last 500 trials of discrimination learning to capture late phase learning, with animals averaging 118.98±3.89 committed trials per day. There was a marginally-significant effect of group (GLM: βgroup =-49.13, p=0.05; **Figure 2-4E**), with EtOH-drinking animals exhibiting more initiation omissions than H2O-drinking animals, but no sex differences (GLM: βsex = 92.88, p=0.21; **Figure 2-4E**) or group*sex interaction (GLM: βgroup*sex = -67.50, p=0.38; **Figure 2-4E**). There were no group (GLM: βgroup =-0.92, p=0.24; **Figures 2-4F, 2-11C**), or sex differences (GLM: βsex =1.26, p=0.20; **Figures 2-4F, 2-11G**) on initiation latencies, as well as no group*sex interaction (GLM: βgroup*sex = -0.76, p=0.51; **Figure 2-4F**). We did, however, find a significant effect of group on incorrect choice latencies (GLM: βgroup = 0.28, p=0.02), with H2O-drinking animals displaying longer latencies than EtOH-drinking animals, but no significant effect of sex, or group*sex interaction was found for this measure. Similar to what was found during early discrimination, there was a significant group effect on choice omissions in late discrimination (GLM: βgroup =1.13, p=0.01), with the H2O-drinking animals exhibiting more choice omissions than the EtOH-

drinking animals, but no effect of sex, or group*sex interaction was found for this measure. However, this effect was largely driven by the same animal as during early discrimination, which upon removal, yielded non-significant results (GLM: βgroup =0.43, p=0.06). Finally, our results indicated that H2O-drinking animals displayed longer reward collection latencies than EtOH-drinking animals (GLM: βgroup =0.23, p=0.01; **Figures 2-11D**), but no effect of sex (**Figure 2-11H**) or group*sex interaction emerged. Thus, the pattern of prior EtOH experience rendering animals more likely to fail to initiate trials was also observed through late discrimination learning, but rats did not take significantly longer to initiate trials when they did so, as in the pretraining phase (which was a trend in early discrimination). Although we did observe greater choice omissions by the H2O-drinking animals during both early and late discrimination, this was driven largely by one animal.

PRL reversal (R) performance

There were no significant group differences (GLM: βgroup = 5.75, p=0.18), sex differences (GLM: βsex =-5.25, p=0.22), or group*sex interaction (GLM: βgroup*sex =-1.88, p=0.74) on sessions to reach criterion for the reversal learning (**Figure 2-5A**). There was an overall average of 25.27±1.62 days to successfully complete the PRL phase.

A GLM model was used to test the effect of drinking group (EtOH, H2O), sex (female, male), days, and their 2-way and 3-way interactions on probability of choosing the better option, number of rewards, and initiation omissions across 15 days of reversal learning. Two females were excluded because they failed to meet criterion for discrimination learning, and never advanced to reversal learning. All other animals demonstrated learning by exhibiting an increase in choosing the better options across days, (GLM: βday =0.01, p=0.01; **Figure 2-5B**), irrespective of group or

sex. There was no effect of group (GLM: βgroup = -0.07, p=0.10), or sex (GLM: βsex =-0.04, p=0.51), and no group*sex (GLM: βgroup* sex = 0.10, p=0.17), sex*day (GLM: βsex* day =0.002, p=0.60), group*day (GLM: βgroup*day =0.002, p=0.47), or group*sex*day (GLM: βgroup*sex* day =-0.004, p=0.46) interactions on the probability of choosing the better option.

All animals increased the number of rewards collected over days, (GLM: βday =0.82, p=0.01; **Figure 2-8C**). There was no significant effect of group (GLM: βgroup=3.69, p=0.75), or sex (GLM: βsex =2.40, p=0.85), and no significant group*sex (GLM: βgroup*sex =5.73, p=0.74), group*day (GLM: βgroup*day=-0.02, p=0.95), sex*day (GLM: βsex*day=0.66, p=0.08), or group*sex*day (GLM: βgroup*sex*day=0.48, p=0.40) interactions on number of rewards collected. All animals decreased the number of initiation omissions across days, (GLM: βday =0.10, p=0.55; **Figure 2-8D**), regardless of group or sex. There was was no significant effect of group (GLM: βgroup= -0.44, p=0.91), or sex (GLM: βsex =-1.81, p=0.61), and no significant group*sex (GLM: βgroup*sex =0.67, p=0.91), group*day (GLM: βgroup*day=-0.06, p=0.78), sex*day (GLM: βsex*day=0.09, p=0.70), or group*sex*day (GLM: βgroup*sex*day=-0.51, p=0.09) interactions on initiation omissions.

Finally, we conducted GLM analyses on the sum of initiation and choice omissions and median (initiation, correct/incorrect choice, reward) latencies in reversal learning, collapsed across all days of testing. Similar to the above measure for across- day learning, we found no significant effect of group (GLM: βgroup =7.88, p=0.97), sex (GLM: βsex =-234.25, p=0.17), or group*sex (GLM: βgroup*sex =-122.1, p=0.59) interaction for initiation omissions. However, though there was no effect of sex (GLM: βsex=0.67, p=0.49), we found that H2O-drinking animals exhibited more choice omissions than EtOH-drinking animals (GLM: βgroup =2.54, p=0.03), with a significant group*sex interaction revealing H2O-drinking females exhibited more choice

omissions than EtOH-drinking females and H2O-drinking males (GLM: βgroup*sex=-3.29, p=0.03). Additionally, there was a marginally-significant effect of group on incorrect choice latencies (GLM: βgroup=0.15, p=0.05), with H2O-drinking animals exhibiting longer incorrect choice latencies than EtOH-drinking animals. Analyses on initiation, incorrect, and reward latencies yielded non-significant results. Unlike for the discrimination phases, upon removal of 2 outliers (> 2 SD from the mean), the effect of drinking group remained significant, with H2O-drinking animals exhibiting more choice omissions than EtOH-drinking animals (GLM: βgroup=1.59, p=0.03).

First 500 trials. An analysis of initiation latencies and omissions, choice (correct and incorrect) latencies and omissions, and reward latencies was conducted for the first 500 trials of reversal learning, to capture early reversal learning. Animals averaged 119.70±8.32 committed trials per day. There were no significant group (GLM: βgroup =-38.04, p=0.65; **Figure 2-5C**), or sex differences (GLM: βsex =-60.29, p=0.41; **Figure 2-5C**) on initiation omissions, and no significant group*sex interaction (GLM: βgroup*sex = 7.67, p=0.94; **Figure 2-5C**). Similarly, there was no effect of drinking group (**Figures 2-5D, 2-12A**), sex (**Figures 2-5D and 2-12E**), or group*sex interaction (**Figure 2-5D**) on initiation latencies, choice omissions, correct choice latencies, or reward collection latencies (**Figures 2-12B, 2-12F**), with the exception of a significant group difference on incorrect choice latencies (GLM: βgroup =0.32, p=0.004), with H2O-drinking animals exhibiting longer latencies when choosing the incorrect stimulus compared to EtOH-drinking animals. There was also a significant group*sex interaction (GLM: βgroup*sex =-0.31, p=0.04) on incorrect choice latency, with H2O-drinking females displaying longer latencies than both H2O-drinking males (p=0.05) and their EtOH-drinking counterparts (p=0.02). In summary, the data for early reversal learning suggests there was no longer any effect of prior EtOH

experience on initiating trials, as had been previously observed during early discrimination learning, but it is important to note that EtOH-drinking animals were less tentative than the H2O-drinking animals given their faster incorrect choice latencies.

Last 500 trials. An analysis of initiation latencies and omissions, choice (correct and incorrect) latencies and omissions, and reward latencies was also conducted for the last 500 trials of reversal learning, to capture late phase learning. Animals averaged 155.11±4.06 committed trials per day. Our analyses on initiation omissions indicated females exhibited more initiation omissions than males (GLM: βsex =-67.96, p=0.02; **Figure 2-5E**), but there was no effect of group (GLM: βgroup =-8.46, p=0.85; **Figure 2-5E**), or group*sex interaction (GLM: βgroup*sex =-8.79, p=0.85; **Figure 2-5E**). There was a marginally-significant effect of sex (GLM: βsex = -2.08, p=0.05; **Figures 2-5F, 2-12G**), with females taking longer to initiate trials compared to males, but no significant group differences (GLM: βgroup = -0.78, p=0.47; **Figures 2-5F, 2-12C**), or group*sex interaction (GLM: βgroup*sex =0.01, p=0.99; **Figure 2-5F**). Females also exhibited more choice omissions than males (GLM: βsex = -0.33, p=0.02), but no group differences or group*sex interaction were found. Upon removal of a single outlier, the effect of sex remained (GLM: βsex = -0.33, p=0.01). Males displayed longer correct choice latencies than females (GLM: βsex = 25.46, p=0.003), whereas H2O-drinking animals exhibited longer reward latencies than EtOH-drinking animals (GLM: βgroup = 0.27, p=0.01; **Figure 2-12D**). We did not find any other significant effect of group, sex, or group*sex interaction.

Thus, the pattern of prior EtOH experience rendering animals more likely to fail to initiate trials and taking longer to do so in early discrimination learning, was not preserved through reversal learning. In summary, the late reversal phase was characterized by predominantly female-

specific attenuations in initiation of trials (both omissions and latencies), as well as correct choice latencies (where males took longer).

Win-Stay/Lose-Shift (WSLS) Strategies

Potential differences in win-stay/lose-shift (WSLS) strategies on stimulus responses employed by each group (EtOH, H2O) and by sex (male, female) were tested for the first 500 committed trials and last 500 committed trials of the discrimination phase and reversal phase by calculating the frequency of each strategy individually. We compared the frequency of using advantageous strategies (i.e., win-stay, lose-shift) vs. less advantageous strategies (i.e., win-shift, lose-stay) by generating an adaptive score: (win-stay + lose-shift) - (win-shift + lose-stay).

There were no significant effects of group, sex, or a group*sex interaction in stimulus-based WSLS strategies individually, or as an 'adaptive score' comparison during early or late discrimination learning (**Figures 2-6A, 2-6B**). However, during early reversal learning, we found a greater use of the lose-shift strategy in the EtOH-drinking rats than the H20-drinking rats (GLM: βgroup =-0.10, p=0.01; **Figures 2-9B**), but a greater use of the less adaptive strategies (GLM: βgroup =-0.20, p=0.0002; **Figure 2-6C)** among the H2O-drinking animals compared to the EtOH-drinking animals. For late reversal learning, we found group differences for the two main stimulus-based strategies (i.e., win-stay and lose-shift), with the H2O-drinking animals using the win-stay strategy (GLM: βgroup = 0.10, p=0.03; **Figure 2-7A**) more than the EtOH-drinking animals, suggesting overall more stimulus persistence in the controls. Conversely, the EtOH-drinking animals used the lose-shift strategies (GLM: βgroup =-0.05, p=0.04; **Figure 2-7B**) more than the H2O-drinking animals, indicating more of an exploration-based strategy. No significant effects were uncovered for adaptive score in the late reversal phase.

DISCUSSION

The present study used an intermittent access model to study the effect of voluntary EtOH consumption on behavioral flexibility using a probabilistic reversal learning paradigm. We included sex as an a-priori moderator. Although forced-exposure models such as intraperitoneal (i.p.) injections (Badanich et al., 2016; Knapp et al., 2012) and EtOH vapor inhalation (Badanich et al., 2011; Willhelm et al., 2015; Planeta et al., 2013) are well-established methods in rodents, they may not be as representative of human alcohol consumption. Therefore, we used a two-bottle choice procedure that allows for oral consumption of EtOH, resulting in increased ecological validity and variability in consumption patterns, which may be important in generating individual differences in alcohol consumption to study subsequent flexible reward learning. To our knowledge, there had only been one study that previously used this voluntary consumption model to test the effects of EtOH on reversal learning, and found no effect (Fisher et al., 2017). However, it is important to note that although the rats in that EtOH group had access to EtOH for 6 weeks, the rats in that study did not demonstrate escalation normally seen with intermittent voluntary EtOH consumption models, including our study. Though we also corroborate no pronounced effects of EtOH exposure on overall learning, a more fine-grained analysis of trial-by-trial and latency data revealed that EtOH-experienced animals were less likely to initiate trials and were slower to initiate trials throughout pretraining and discrimination learning. As mentioned in the Introduction, a failure to initiate trials and taking longer to do so (together with intact stimulus-response and reward collection times), points to a deficit in attention to task (i.e. task engagement) following EtOH experience, not a problem with motivation to learn about stimuli that predict rewards and to procure the rewards themselves. Collectively, the data support the interpretation that the most pronounced attentional decrements appear closest in time to drinking, despite intact

motivation for reward and motor responding throughout learning. We further elaborate on these attentional effects, as well as reversal-specific EtOH effects on WSLS strategies below.

Consumption patterns of EtOH

We observed an escalation of drinking over the course of the twenty-nine days of alcohol access, irrespective of sex, an expected pattern when using intermittent-access models compared to continuous access models. Several studies administering intermittent exposure have shown that alternating brief periods of alcohol access with brief periods of no access can actually escalate alcohol consumption to excessive levels (Hwa et al., 2011; Rosenwasser et al., 2012; Becker et al., 2014; Carnicella et al. 2009, 2014; Simms et al., 2008; George et al., 2012) compared to continuous daily access (Hwa et al., 2011; Rosenwasser et al., 2012; Becker et al., 2014) which typically exhibit more moderate, but stable levels of intake. However, despite the recent popularity of the intermittent-access model, the underlying psychological and neurobiological mechanisms that promote the escalation of alcohol consumption remain unclear and should be investigated in future studies.

We found that females reached a higher EtOH consumption level and exhibited greater high-drinking days than males. These findings are consistent with previous studies showing that female rodents drink more EtOH than males (Vetter-O'Hagen et al., 2009, 2011; Lourdes de la Torre et al., 2015; Hwa et al., 2011) and exhibit less aversion to EtOH, as demonstrated by conditioned taste aversion using EtOH-saccharin pairings (Cailhol et al., 2002; Sherrill et al., 2011; Schramm-Sapyta et al., 2014), with males developing an aversion after only one pairing and females after the third pairing and only at higher doses of EtOH (Sherrill et al., 2011). Other groups have reported that the rewarding effects of EtOH are enhanced in females and therefore, may be

hormone-dependent (Torres et al., 2014), which may explain the increased EtOH intake over time that may lead to increases in potential for over-consumption. However, the role of gonadal hormones on ethanol intake and preference remain unclear, as other studies have shown that the removal of testicular hormones in males decreases alcohol intake, and no differential consumption in ovariectomized vs. intact females (Vetter-O'Hagen et al., 2011). Although seemingly contradictory, these findings may provide further evidence of the dissociation between chromosomal and gonadal sex, given that studies have found alcohol reinforcement is mediated by chromosomal sex, independent of gonadal phenotype (Barker et al., 2010). Taken together, the present findings add to a growing body of evidence for sex differences in alcohol consumption patterns.

Attentional deficits following EtOH across learning stages

We observed the most pronounced impairment following EtOH on sessions to reach criterion during pretraining; a pattern that was not maintained through discrimination or reversal learning. Prior studies testing the relationship between alcohol exposure and performance on reversal learning tasks have largely been mixed, with some studies demonstrating alcohol produced impairments in both discrimination and reversal learning (Fernandez et al., 2016; Kuzmin et al., 2012), other showing no impairments for either (Fernandez et al., 2017; Fisher et al., 2017; Badanich et al., 2016; DePoy et al., 2013; Kroener et al., 2012), and some only showing impairments on reversal, with the discrimination learning phase largely intact (Fernandez et al., 2017; Badanich et al., 2011; Kuzmin et al., 2012; Coleman et al., 2011; Obernier et al., 2002). These conflicting findings may be due to variations in alcohol administration procedures, most of which have used forced-exposure models, variations in maximum blood ethanol concentration

(BEC) levels, and/or types of reversal learning paradigms employed. Even similar methods of administrations produce variable results with studies showing vapor EtOH exposure impairs reversal learning (Badanich et al.,2011), does not impair (Kroener et al., 2012), or improves reversal learning (DePoy et al., 2013), with the doses of i.p. EtOH administration determining whether an impairment is observed (Badanich et al., 2016). Our study is consistent with the only other study to our knowledge that has used a 2-bottle choice procedure to assess effects on reversal learning, which similarly found no overt learning impairment (Fisher et al., 2017). Although we did not measure BEC levels during the 2-bottle choice procedure, it is likely they never reached BECs known to impair reversal learning based on previous experiments using forced-exposure models (150-550 mg/dl). Task parameters (i.e. stimulus modalities, probability of reward) may also contribute to differential effects, with groups using lever or touchscreen-based responding reporting no pronounced impairments on discrimination or reversal learning (Fernandez et al., 2017; Fisher et al., 2017; DePoy et al., 2013), whereas groups employing Morris Water and Barnes Maze tasks reporting impairments in reversal learning. Discrimination learning seems to remain mostly intact across diverse paradigms (Kuzmin et al., 2012; Coleman et al., 2011, 2014; Obernier et al., 2002). We maintain that many past studies of EtOH effects on reversal learning typically report omnibus measures of learning and do not probe more micro (trial-by-trial) analyses that may be more sensitive to EtOH effects, as reported here.

We found impairments associated with prior EtOH experience, such that animals previously exposed to EtOH required more sessions to reach criterion and exhibited longer initiation and choice latencies during pretraining. Similar deficits were also found during discrimination learning, with a greater number of initiation omissions and longer initiation latencies in the EtOH-experienced group. These effects were most pronounced in female animals

(discussed below). There have been studies of EtOH exposure on attention using the 5-choice serial reaction time task (Sanchez-Roige et al., 2014a, 2014b; Irimia et al., 2014, 2015; Boutros et al., 2017, Brys et al., 2014; Givens, 1997; Louth et al., 2016; Slawecki et al., 2006), considered to be the gold standard for measuring attention in rodents, revealing that EtOH-exposed animals exhibit attention deficits. An evaluation of attentional capacity using a 5-choice continuous performance task following EtOH exposure in rats found this group exhibited more omissions and longer choice latencies relative to control rats, while motivation remained intact. Indeed, there were no differences in accuracy or reward latencies (Irimia et al., 2014), similar to our present findings. However, it is important to note these differences were observed only during acute, not prolonged, abstinence from EtOH exposure- the latter, as we report here. Other groups have previously reported an EtOH dose-dependent decrease in the ability to direct and sustain attention to brief stimuli, but not a complete disruption in overall performance (i.e. percentage correct), also suggestive of an impairment in attentional processing (Givens, 1997). Similarly, we found differences in measures of attention processing (i.e. initiation omissions and latencies), but observed no overall performance deficit in the probability of choosing the better option for both the discrimination and reversal phases of learning. Importantly, attentional deficits following EtOH experience have also been found in human binge-drinkers (i.e. more omitted trials, lower accuracy), particularly under task variants meant to increase attentional load in a human version of the 5-CSRTT (Sanchez-Roige et al., 2014a). Indeed, the pattern we observed here- that EtOH-drinking animals exhibited more initiation omissions and longer initiation latencies (particularly in early phases of pretraining and discrimination)- stand in contrast to their quick reward collection times and intact accuracy measures, relative to H2O-drinking animals in these same phases. It is, however, possible that trial initiations are simply more sensitive measures of motivation and more

easily perturbed following EtOH than reward collection or stimulus-response times, but given the convergence of evidence outlined above this is unlikely.

Sex differences in pretraining and reversal learning

Some interesting sex differences on reversal learning emerged in our experiment. The most pronounced impairments were observed during late reversal, with females exhibiting greater omissions (initiation and choice) and longer initiation latencies than males, irrespective of prior EtOH exposure. This is in agreement with the human literature, which has shown that males outperform females on reversal learning (Overman, 2004; Evans and Hampson, 2015), and with observations in marmosets where females require more trials to learn reversals than males. Interestingly, though we find sex differences in reversal learning, there were no differences in the number of omitted trials or reaction times (i.e. latencies) (LaClair et al., 2019). Relevant to this, Grissom et al. (2019) conducted an extensive review of sex differences in several aspects of executive function, including attention, and did not find evidence to support robust sex differences in this domain. Prior studies have reported that male rodents show higher levels of novelty-seeking (Palanze et al., 2001), with higher novelty-seeking related to higher levels of impulsivity in males relative to females (Lukkes et al., 2016).

Sex-dependent EtOH effects in early learning and attentional measures

A sex-dependent drinking group difference was observed, with EtOH-exposed females more affected than males on measures of attention: they exhibited more initiation omissions than their H2O-drinking counterparts during both pretraining and early discrimination learning, which is also reflected in a greater number of sessions required to reach criterion in early learning.

Although prior research has not provided sufficient evidence supporting an attentional deficit specific to females (Grissom and Reyes, 2019), there is now substantial evidence to support a potential EtOH-specific effect on attentional processing (Sanchez-Roige et al., 2014a; Irimia et al., 2014; Givens, 1997). Therefore, it is plausible that sex effects we observe here are moderated by EtOH-experience, resulting in more pronounced deficits in attentional processing in EtOH-drinking females. It is worth noting that EtOH-exposed males also exhibited some impairments (i.e. more sessions to reach criterion and initiation omissions), but this effect was only observed early in pretraining and did not extend to discrimination or reversal learning.

Sex-independent EtOH effects on WSLS in reversal learning

We observed sex-independent EtOH effects on WSLS strategies during reversal learning. Rats with prior EtOH experience were more likely to use a "shift" strategy whereas H2O-drinking animals were more likely to "stay" with the previous stimulus choice in the reversal phase. Similarly, animals with prior EtOH experience were generally more flexible in early reversal learning (i.e. they exhibited a greater 'adaptive' score) than H2O-drinking animals. This suggests that EtOH-experienced rats had a more tenuous representation of trial-by-trial stimulus-based contingencies upon criterion-level performance than the H2O-drinking control rats, and could consequently be more flexible. However, all rats generally increased their choice of the better option and rewards collected over time, while decreasing the number of initiation omissions for both the discrimination and reversal phase. The lack of pronounced EtOH impairments on overall discrimination and reversal phases of learning- as measured by global measures such as the probability choosing the better option over time- may be attributed to plasticity following protracted abstinence in rodents (Crews and Nixon et al., 2008; Nixon et al., 2008; Nixon and

Crews, 2004), and humans (O'Neill et al., 2001; Rosenbloom et al., 2007). Similarly, the probability of using WSLS strategies across time was ~0.5, suggesting these strategies were not used effectively for learning. Indeed, dissociations in learning and WSLS have been reported before (Verharen et al., 2020). It will be important to investigate the extent to which the later pro-exploratory phenotype relies on an early attentional decrement, or if these are orthogonal effects of chronic EtOH experience.

Conclusions

In summary, we observed pronounced trial initiation omissions following EtOH experience in females during pretraining and discrimination learning. These phases are closest in time to the last EtOH experience and constitute the early abstinence period. Additionally, this attentional decrement, which was most pronounced in female animals, was partnered by an enhanced exploration strategy in all EtOH drinking animals, both males and females, later in reversal learning.

Alterations related to attention and processing speed in early EtOH abstinence (during pretraining) may have a domino effect on later learning, leading to the sex by drinking group interaction we observe in discrimination learning, and perhaps contribute to the enhanced exploration phenotype in reversal learning. A true test of this would require animals to undergo pretraining, discrimination learning, and drinking prior to any reversal learning. Ultimately, all rats exhibited intact motivation and motor timing, and were able to increase their probability of choosing the better option and number of rewards, while decreasing their initiation omissions. Although voluntary alcohol consumption models, such as the one employed here, do not model severe alcohol dependence like forced-exposure models, they do however reflect escalating,

chronic intermittent drinking that corresponds to the early stages of problematic drinking, before individuals transition to alcohol dependence. Attenuated attentional mechanisms in early abstinence may not contribute to decrements in flexible learning per se, but may instead detract from executive functions important in limiting (over)consumption. Future studies should investigate the brain mechanisms and the role of gonadal hormones on alcohol consumption and attention, and systematically compare these measures as predictors of consumption (i.e. relapse) during acute vs. prolonged abstinence.

METHODS

A timeline of all procedures is shown in **Figure 2-1A**. Subjects were adult male (n=16) and female (n=16) Long-Evans rats (Charles River Laboratories). All animals were between postnatal day (PND) 60-70 upon start of EtOH or H2O-only consumption and between PND 130-140 at the start of behavioral testing. All rats underwent a 3 day acclimation period, during which they were pair-housed and given food and water ad libitum, and remained in cages with no investigator interference. Following the 3-day acclimation period, animals were handled for 10 min per animal for 5 consecutive days. During the handling period, the animals were given unlimited food and water access and were tail marked. After the handling period, animals were singly-housed under standard housing conditions (room temperature 22–24° C) with a standard 12 h light/dark cycle (lights on at 6am). This study was conducted in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Chancellor's Animal Research Committee at the University of California, Los Angeles.

Rodent voluntary alcohol regimen: 2-bottle choice procedure

Rat home cages were modified to allow for the placement of 2 bottles. Rats (n=16; n=8 male, n=8 female) were given access to both water and 20% alcohol simultaneously, with placement of bottles counterbalanced, for a 24-hour period 3 days per week, and only water on the remaining days. Alcohol access was terminated at 10 weeks (29 days of access) after animals' EtOH consumption stopped escalating. An age-matched control cohort of water (H2O)-only drinking animals (n=16; n=8 male, n=8 female) was placed in modified home cages allowing for the placement of 2 bottles with water-only also for a total of 10 weeks. Weight of bottles was measured before and after alcohol and/or water-only access to measure daily consumption amounts with a control cage placed on the same rack to account for leakage.

Behavioral task: rodent probabilistic reversal learning (PRL) task

Immediately following the termination of the consumption period, all animals were placed on food restriction to 14 grams/day (females) or 18 grams/day (males) of chow for 5 days prior to behavioral testing. Animals were weighed every other day and monitored closely to not fall below 85% of their maximum, free-feeding weight. Behavioral testing was conducted in operant conditioning chambers outfitted with an LCD touchscreen opposing the sucrose pellet dispenser (Stolyarova and Izquierdo, 2017; Stolyarova and Izquierdo, 2015) during the animals' inactive phase. Rewards were 45mg sucrose pellets (Dustless Precision Pellets #F0023, Bio-Serv). All chamber equipment was controlled by customized ABET II TOUCH software. Following 5 days of forced abstinence (n=16; EtOH group) or a rest period (n=16; H2O group), animals began pretraining.

The pretraining protocol, adapted from established procedures described in Stolyarova et al. (2017) (Stolyarova and Izquierdo, 2017), consisted of a series of stages: Habituation, Initiation Touch to Center Training (ITCT), Immediate Reward Training (IMT), designed to train rats to nosepoke, initiate a trial, and select a stimulus to obtain reward. During habituation, rats were required to eat five sucrose pellets out of the pellet dispenser inside the chambers within 15 min before exposure to any stimuli on the touchscreen. ITCT began with the display of white graphic stimuli on the black background of the touchscreen. During this stage, a trial could be terminated for one of two reasons: if a rat touched the displayed image (and received reward), or if the image display time (40 s) ended, after which the stimulus disappeared, a black background was displayed, and a 10 s inter-trial interval (ITI) ensued. If the rat did not touch within 40 s this was scored as an initiation omission. IMT began in the same way as ITCT, but the disappearance of the white graphic stimulus was now paired with the onset of a target image immediately to the left or right of the stimulus (i.e. forced-choice) that the rat was required to nosepoke to obtain reward. During this stage, a trial could be terminated for one of three reasons. First, if a rat touched the center display (i.e. white graphic stimulus) and touched the image displayed on either side, after which there was a dispensation of one sucrose pellet and illumination of the tray-light. Second, if the rat failed to touch the center white graphic stimulus before the display time ended (40 s), the stimulus disappeared, a black background was displayed, and a 10 s ITI ensued, scored as a initiation omission. Third, if the image display time (60 s) ended, after which the stimulus disappeared, a black background was displayed, and a 10 s ITI ensued, scored as a choice omission. For habituation pretraining, the criterion for advancement was collection of all 5 sucrose pellets. For ITCT, the criterion to the next stage was set to 60 rewards consumed in 45 min. The criterion for IMT was set to 60 rewards consumed in 45 min across two consecutive days.

After completion of all pretraining schedules, rats were advanced to the discrimination phase of the PRL task, in which they would initiate a trial by touching the white graphic stimulus in the center screen (displayed for 40 s), and choose between two visual stimuli presented on the left and right side of the screen (displayed for 60 s) counterbalanced between trials, assigned as the Better or Worse options, rewarded with a sucrose pellet, with probability pR(B)=0.70 and pR(W)=0.30, respectively. Assignment of the stimulus to better or worse reinforcement was counterbalanced across conditions. If a trial was not initiated within 40 s, it was scored as an initiation omission. If a stimulus was not chosen, it was scored as a choice omission, and a 10 s ITI ensued. If a trial was not rewarded, a 5 s time-out would follow, subsequently followed by a 10 s ITI. Finally, if a trial was rewarded, a 10 s ITI would follow after the reward was collected (**Figure 2-1B**). The criterion was set to 60 or more rewards consumed and selection of the better option in 80% of the trials or higher during a 60 min session across two consecutive days. After reaching criterion for the discrimination phase, the rats advanced to the reversal phase beginning on the next session. During the reversal phase, rats were required to remap stimulus-reward contingencies. The criterion for the reversal phase was the same as the discrimination phase.

Statistical analyses

To test the study hypotheses, a series of mixed-effects General Linear Models (GLM) and ANOVA analyses were conducted using MATLAB (MathWorks, Natick, Massachusetts; Version R2018b) (MATLAB, 2013) and SPSS (IBM SPSS Statistics, Version 25) (IBM SPSS, 2017). MATLAB was also used for graphing.

The EtOH consumption data were analyzed with ANOVAs with sex and drinking-group as between-subject factors, and EtOH consumption days (D1-D29) as a within-subject factor, with

EtOH consumption (g/kg) as the primary outcome (Morales et al., 2015; Priddy et al., 2017; Piano et al., 2005). Independent samples t-tests were conducted for EtOH-drinking related varibles, such as maximum EtOH consumption, calculated as the highest amount of daily EtOH consumption reached over the course of the 29 days of alcohol access averaged by sex, and the number of high-drinking days (i.e. 5+ g/kg/24 hrs) (Leeman et al., 2010).

Learning data (sessions to criterion, probability correct), number of rewards collected, omission, and latency data were analyzed with GLM in MATLAB (fitglme function; Statistics and Machine Learning Toolbox; MathWorks, Natick, Massachusetts; Version R2017a). Probability correct, number of rewards, and initiation omissions, were analyzed using GLM across repeated days of testing per animal with drinking group (EtOH vs. H2O) and sex (female vs male) as fixed effects and animal as a random effect. In Figures, we show the first 15-20 days of learning to avoid overweighting performance of increasingly fewer animals at the extremes. Omission (sums) and latency data (medians) included one observation per subject and were also analyzed with a GLM with drinking group and sex as fixed effects and animal as a random effect. All post-hoc tests were Bonferroni-corrected to account for the number of comparisons. Statistical significance was noted when p-values were less than 0.05, p-values between 0.05 and 0.06 are reported as marginally significant.

Trial-by-trial analyses were conducted to investigate potential sex and group differences in WSLS strategies used in the PRL task. WSLS are strategies commonly examined in decision-making tasks involving risk and reward that can reveal changes in sensitivity to surprising outcomes and feedback learning. Each trial was classified as a win if an animal received a sugar pellet, and as a lose trial if no reward was delivered. We classified decisions as Win-Stay when a rat chose the same stimulus on the subsequent trial after a win and as Lose-Shift when the rat

switched to the alternative stimulus after a loss. We also studied less advantageous strategies: we classified selecting a different stimulus after being rewarded as Win-Shift, and choosing the same stimulus after not being rewarded as Lose-Stay. We calculated the frequency of each type of event (win-stay, lose-shift, win-shift, lose-stay). For win-stay events we divided the total number of times an animal chose the same stimulus on the trial following a win, by the total number of wins [sum(win-stay)/sum(win)]. For lose-shift events we divided the total number of times an animal chose the alternative stimulus on a trial following a loss, by total number of lose trials [sum(lose-shift)/sum(lose)]. Win-shifts events were inversely proportional to win-stay events [1-sum(win-stay)]. Likewise, lose-stay events were inversely proportional to lose-shift events [1-sum(lose-shift)]. Finally, we compared the frequency of using advantageous strategies (i.e., win-stay, lose-shift) vs. less advantageous strategies (i.e., win-shift, lose-stay) by generating an adaptive score: (win-stay + lose-shift) - (win-shift + lose-stay). Ultimately, we compared the use of advantageous (i.e., win-stay, lose-shift) with less advantageous (i.e., win-shift, lose-stay) strategies by drinking group and sex.

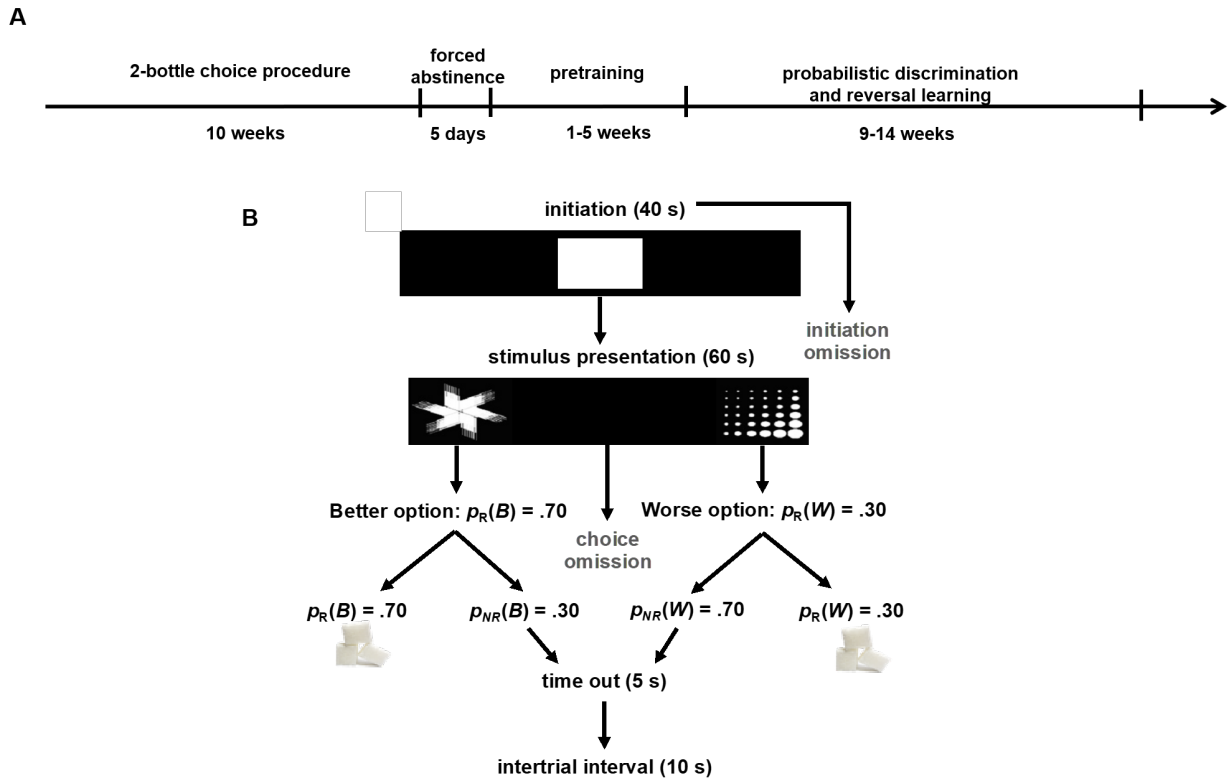

Figure 2-1. Experimental timeline and probabilistic discrimination and reversal learning paradigm. **(A)** Sequence of events during the voluntary EtOH drinking regimen (i.e. 2-bottle choice procedure), forced abstinence, and behavioral testing, depicted from left to right. **(B)** Task structure of probabilistic reversal learning (PRL), in which the animal initiated a trial by nosepoking the white graphic stimulus in the center screen (displayed for 40 s), and chose between two visual stimuli pseudorandomly presented on either the left and right side of the screen (displayed for 60 s), assigned as the Better or Worse options, rewarded with a sucrose pellet with probability $p_R(B)=0.70$ and $p_R(W)=0.30$, respectively. If a trial was not rewarded [$p_{NR}(B)$ or $p_{NR}(W)$], a 5 s time-out would commence. If a stimulus was not chosen, it was considered an omission and a 10 s ITI would commence. If a trial was rewarded, a 10 s ITI would follow reward collection.

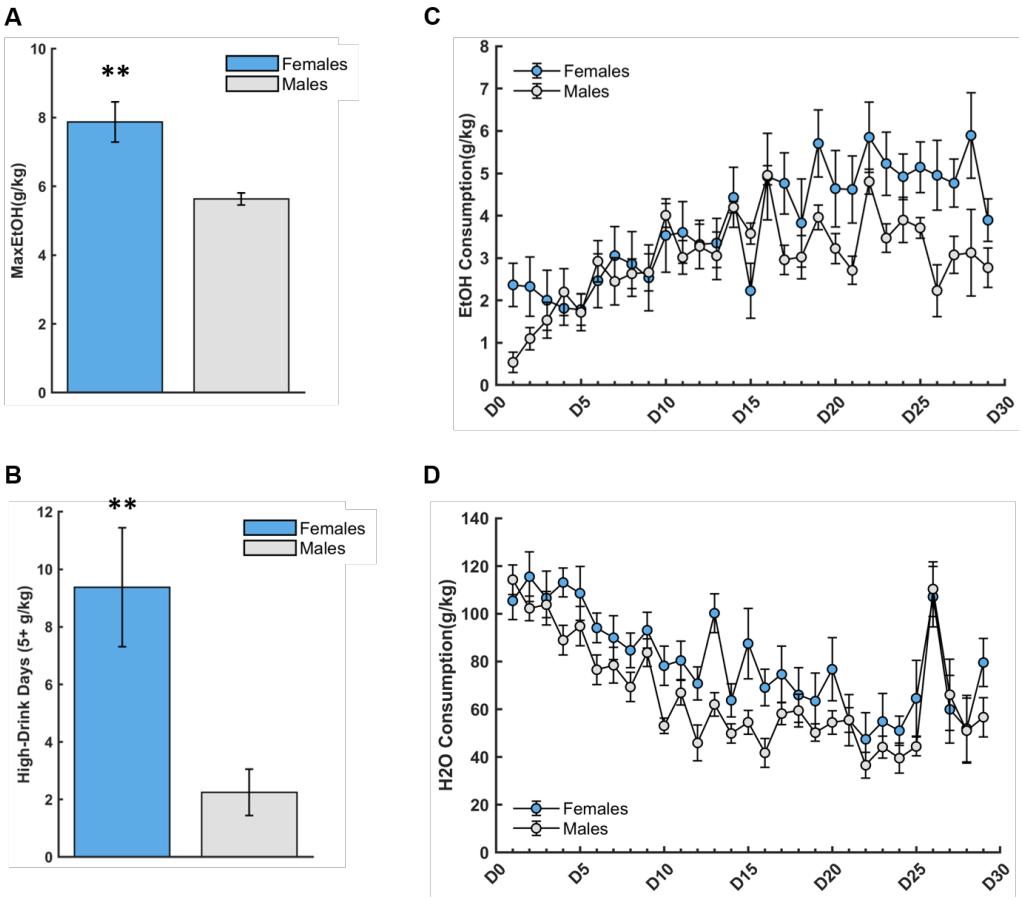

Figure 2-2. 20% Ethanol (EtOH) consumption (g/kg) across drinking days (D1-D29) is more pronounced in females. (**A**) Females reached a greater maximum EtOH level of consumption than males. (**B**) Females exhibited more high-drinking (5-10 g/kg/24 hr) days than males. (**C**) There was a within-subject effect of day, with both males and females escalating their EtOH consumption over time, a marginally significant sex*day interaction, and a trend for a main effect of sex. (**D**) There was a de-escalation of water consumption over the 29 days of alcohol access, but no main effect of sex or sex*day interaction. Bars indicate ± S.E.M. n=8 males, n=8 females, **p<0.01.

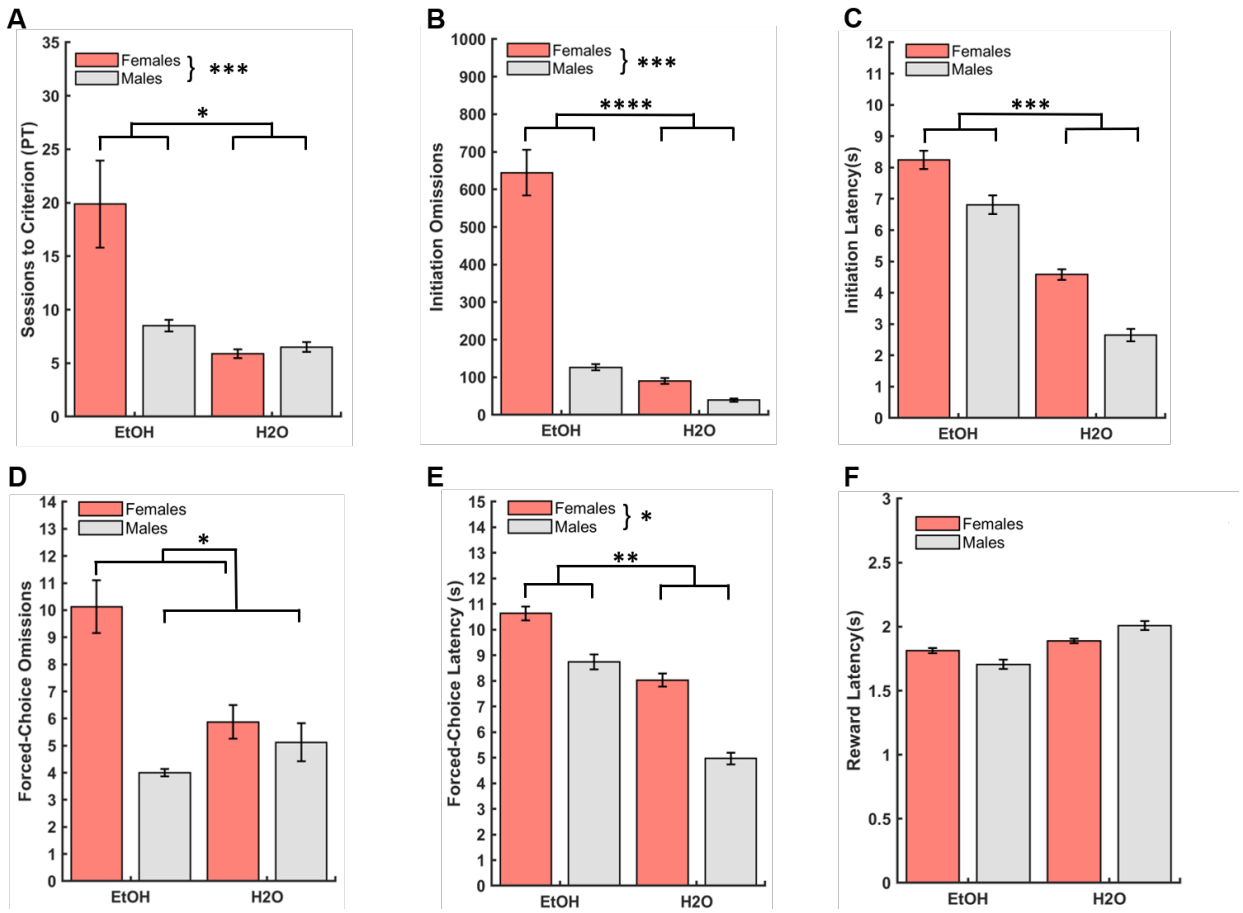

Figure 2-3. Drinking group, sex differences, and drinking group by sex interactions in operant pretraining. (**A**) The number of pretraining sessions required to reach criterion to advance to the main PRL task was greater for EtOH-drinking (male and female) animals than their H2O-only drinking counterparts, and greater for EtOH-drinking females compared to EtOH-drinking males. Although a significant group*sex interaction was found for this measure, signifcant pairwise comparisons are not depicted for clarity. (**B**) EtOH group and females exhibited more initiation omissions than the H2O group and males, respectively. Although a significant group*sex interaction was also found for this measure, significant pairwise comparisons are not depicted for clarity. (**C**) EtOH group exhibited longer initiation latencies than the H2O group. No sex differences were found for initiation omissions. (**D**) Females displayed more forced-choice omissions. No group differences were found for forced-choice omissions. (**E**) EtOH group and females exhibited longer forced-choice latencies than the H2O group and males, respectively. **(F)** No group or sex differences were found for reward latencies. Latencies represent group medians. Bars indicate $\pm S.E.M.$ n=16 males, n=16 females, *p ≤0.05, **p ≤0.01, ***p ≤0.001, ****p≤0.0001

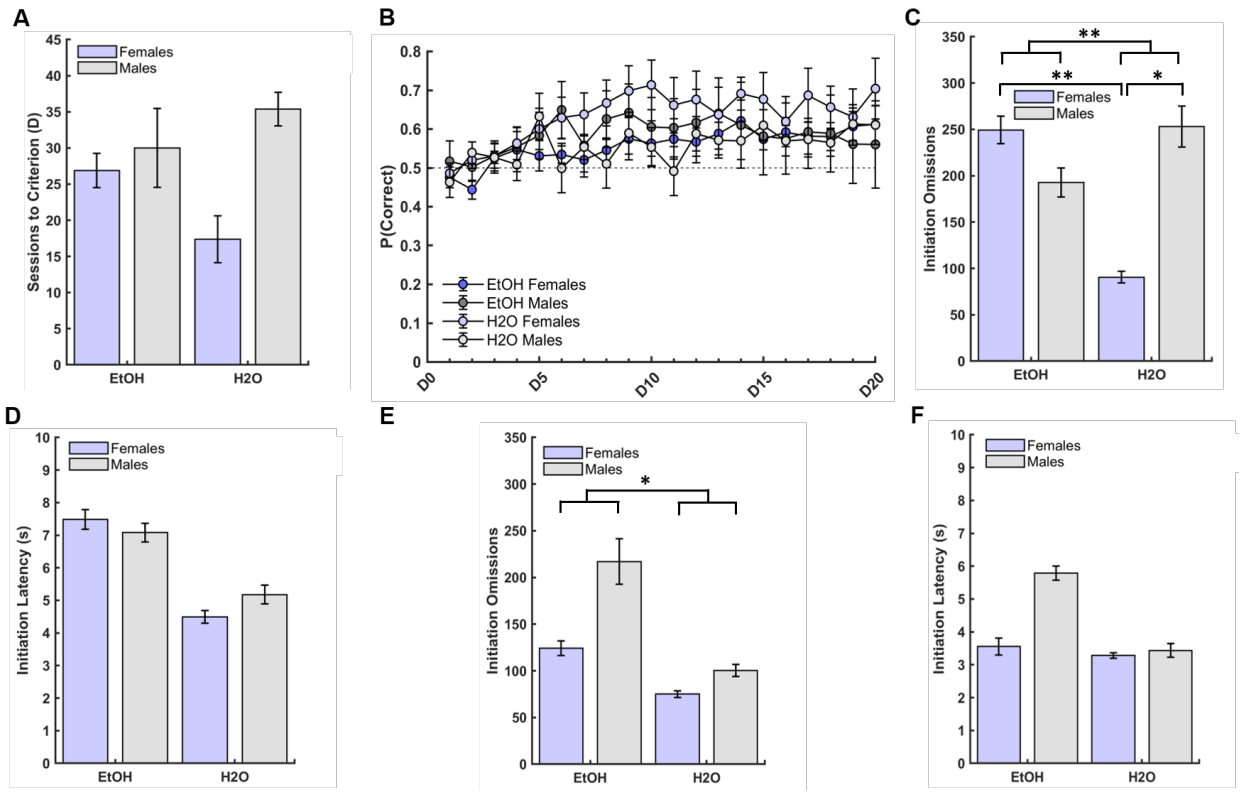

Figure 2-4. Drinking group differences in initation omissions, but not latencies, during early and late probabilistic discrimination learning performance. (**A**) There were no group or sex differences in the number of sessions to reach criterion for the discrimination learning phase, only a marginally significant group*sex interaction. (**B**) There was no effect of group or sex on probability of choosing the better option, only an effect of day (D1-D20), with the probability of choosing the better option increasing across testing days irrespective of sex or group. (**C**) EtOH group exhibited more initiation omissions compared to the H2O group. EtOH-drinking females exhibited more initiation omissions than H2O-drinking females, and H2O-drinking males exhibited marginally more initiation omissions than H2O-drinking females in early discrimination learning. (**D**) No group or sex differences emerged for initiation latencies during early discrimination learning. (**E**) EtOH-drinking animals (male and female) exhibited marginally more initiation omissions than H2O-drinking animals, but no sex differences emerged for late discrimination learning. (**F**) No group or sex differences were found for initiation latencies during late discrimination learning. Latencies were medians. Bars indicate $\pm\, S.E.M.$ n=16 males, n=16 females, *p ≤0.05, **p ≤0.01

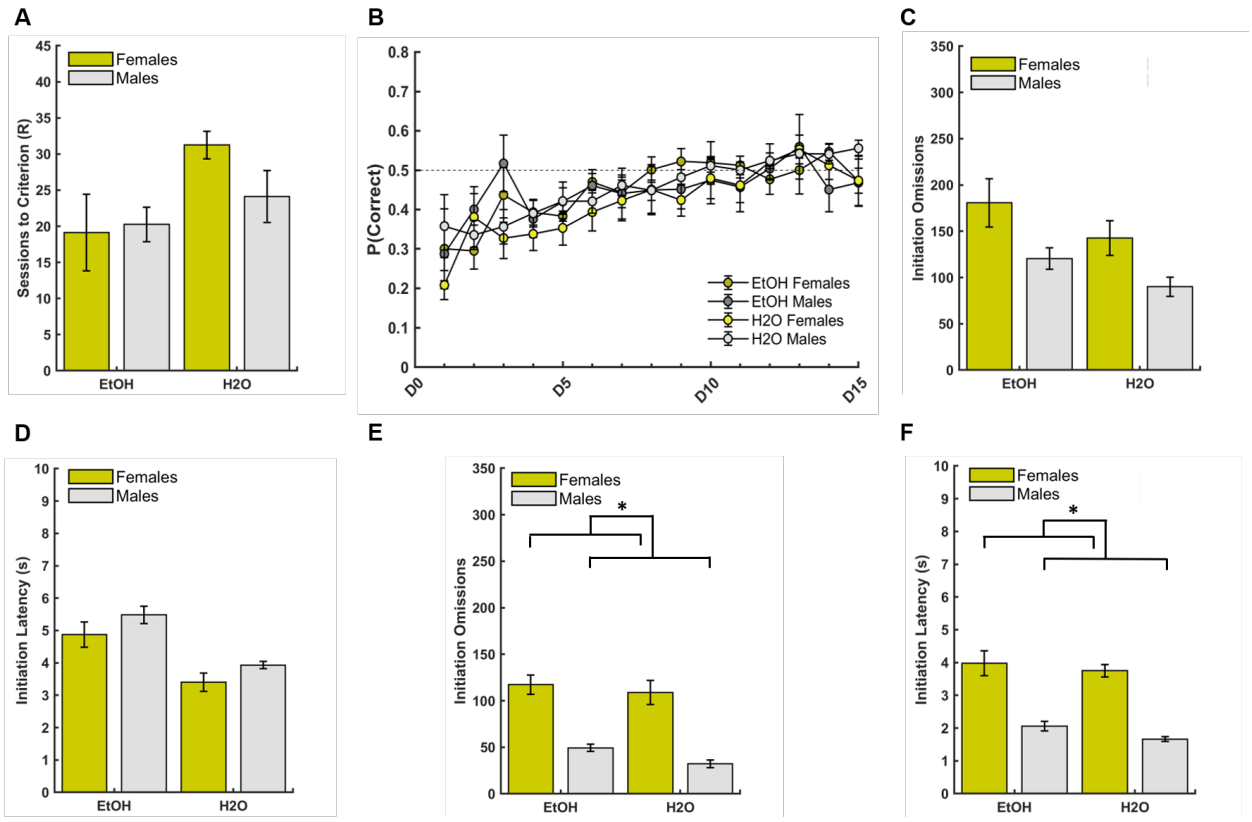

Figure 2-5. Sex differences, but no drinking group differences, in initiation omissions and latencies during late probabilistic reversal learning performance. (**A**) There were no group or sex differences in the number of sessions to reach criterion for the reversal learning phase. (**B**) There was no effect of group or sex on probability of choosing the better option, only an effect of day (D1-D25), with the probability of choosing the better option increasing across the fifteen testing days regardless of sex or group. (**C**) No group or sex differences were found for initiation omissions during early reversal learning. (**D**) No group or sex differences were found for initiation latencies during early reversal learning. (**E**) Females exhibited more initiation omissions than males during late reversal learning. (**F**) Females exhibited marginally longer initiation latencies than males during late reversal learning. Latencies represent group medians. Bars indicate $\pm\, S.E.M.$ n=16 males, n=14 females, *p ≤0.05

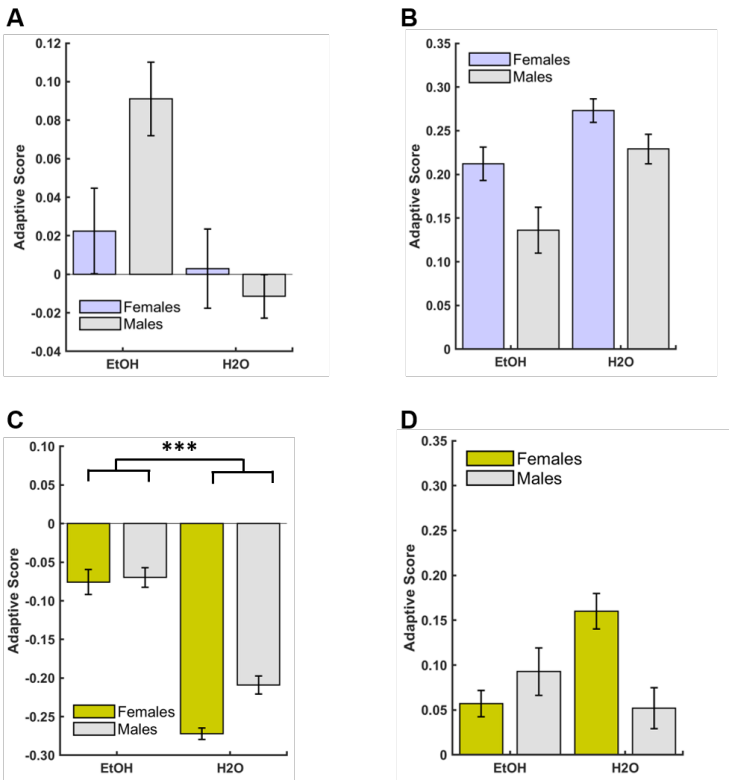

Figure 2-6. Drinking group differences in early reversal learning strategies. An adaptive score was calculated as the difference between advantageous strategies and less advantageous strategies: (win-stay + lose-shift) - (win-shift + lose-stay). (**A**) There were no group or sex differences in adaptive scores during early discrimination learning. (**B**) There were no group or sex differences in adaptive scores during late discrimination learning. (**C**) EtOH group exhibited higher adaptive scores than H2O group during early reversal learning, but there were no sex differences. (**D**) There were no group or sex differences in adaptive scores during late reversal learning. Bars indicate $\pm$ $S.E.M.$ n=16 males, n=16 females, *p ≤0.05, **p ≤0.01 ***p ≤0.001

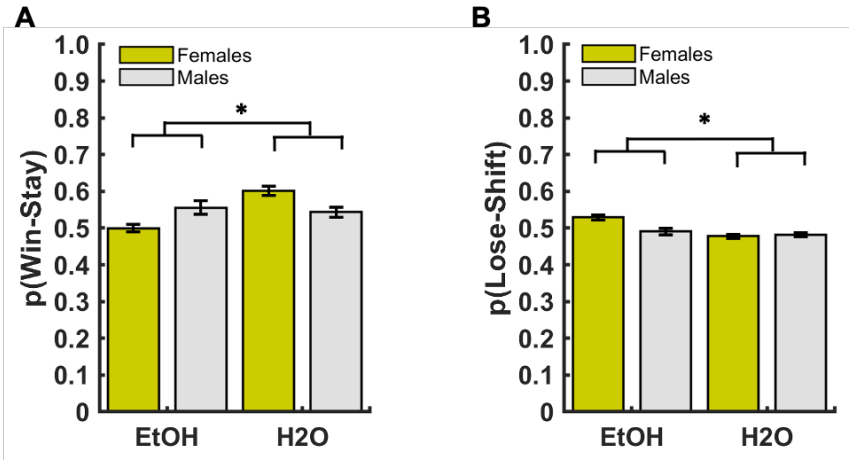

Figure 2-7. Drinking group differences in the use of Win-Stay/Lose-Shift strategies during late probabilistic reversal learning. (**A**) H2O-drinking animals used the win-stay strategy more than EtOH-drinking animals. (**B**) EtOH-drinking animals used the lose-shift strategy more than H2O-drinking animals. Bars indicate $\pm S.E.M.$ *p ≤0.05

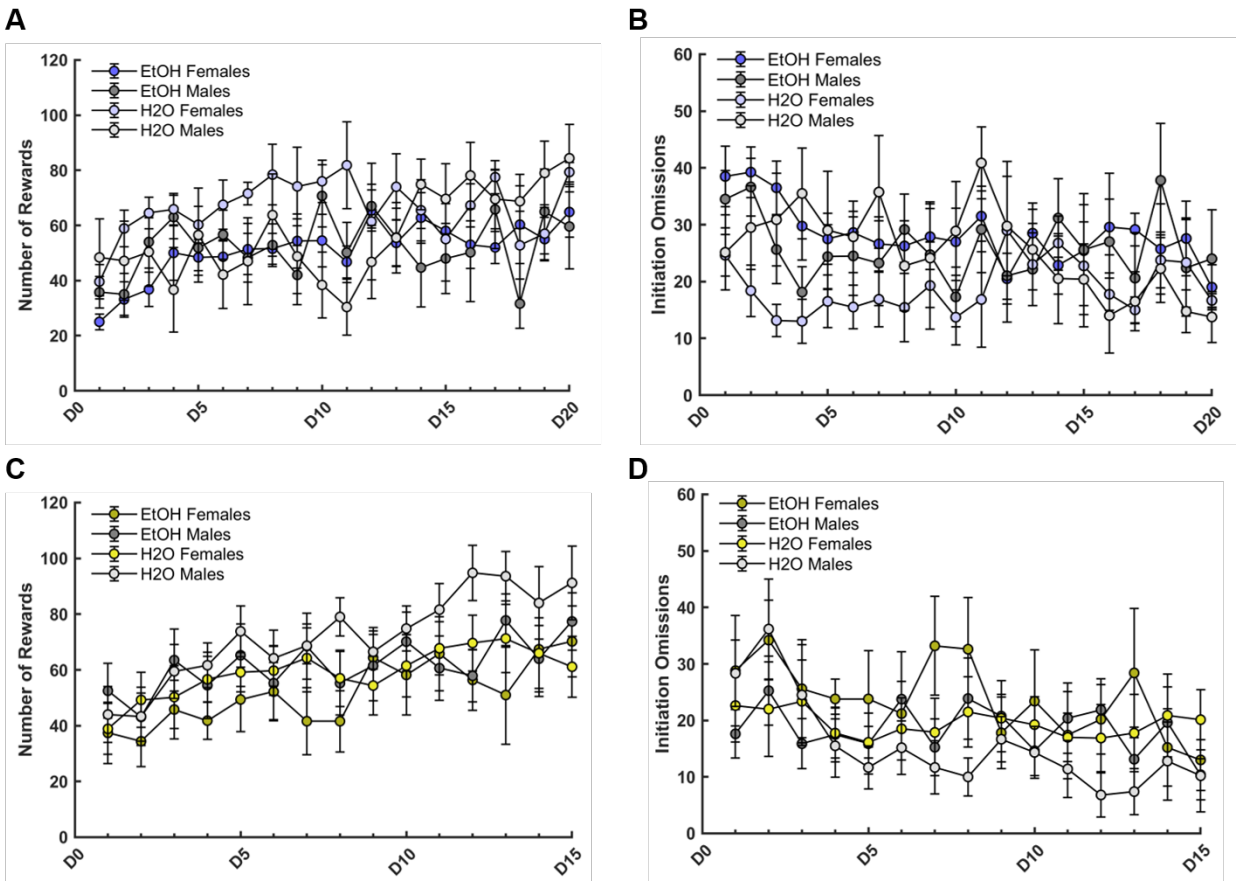

Figure 2-8. Drinking group and sex differences on number of rewards and initiation omissions during probabilistic discrimination and reversal learning. (**A**) All animals regardless of drinking group or sex increased their number of rewards collected over the twenty testing days of discrimination learning. H2O-drinking animals and males collected a greater number of rewards than EtOH-drinking animals and females, respectively. Females displayed a greater increase in rewards collected across days. (**B**) All animals regardless of drinking group or sex decreased their initiation omissions over the twenty testing days of discrimination learning. EtOH-drinking animals and females had more initiation omissions than H2O-drinking animals and males, respectively. EtOH-drinking females displayed more initiation omissions than H2O-drinking females. (**C**) All animals regardless of drinking group or sex increased their number of rewards collected over the fifteen testing days of reversal learning. There were no group or sex differences on number of rewards collected. (**D**) All animals regardless of drinking group or sex decreased the number of initiation omissions over the fifteen testing days of reversal learning. There were no group or sex differences on initiation omissions. Bars indicate $\pm S.E.M.$

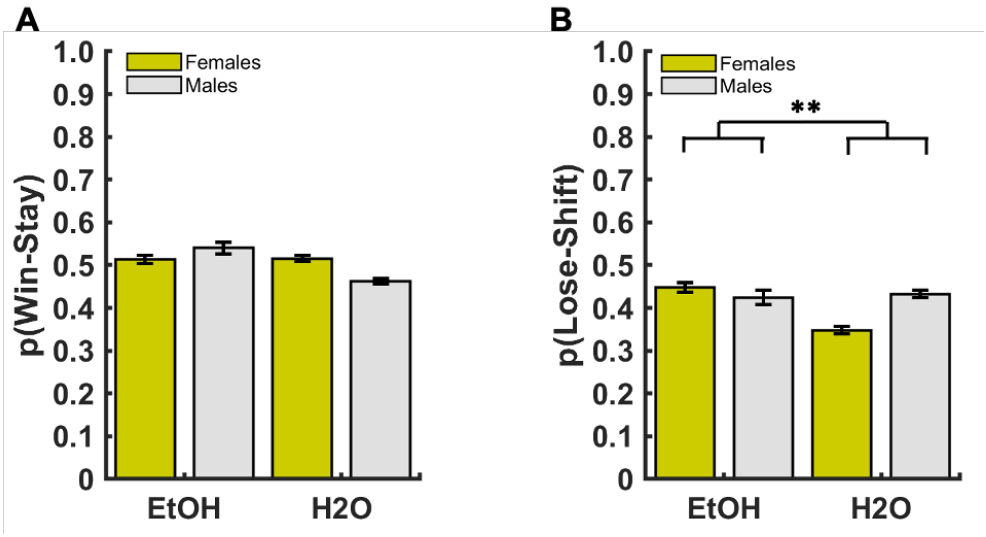

Figure 2-9. Drinking group differences in use of lose-shift and lose-stay strategies during early probabilistic reversal learning. (**A**) No group or sex differences in the use of the win-stay strategy. (**B**) EtOH-drinking animals used the lose-shift strategy more than H2O-drinking animals. Bars indicate $\pm S.E.M.$ **p ≤0.01

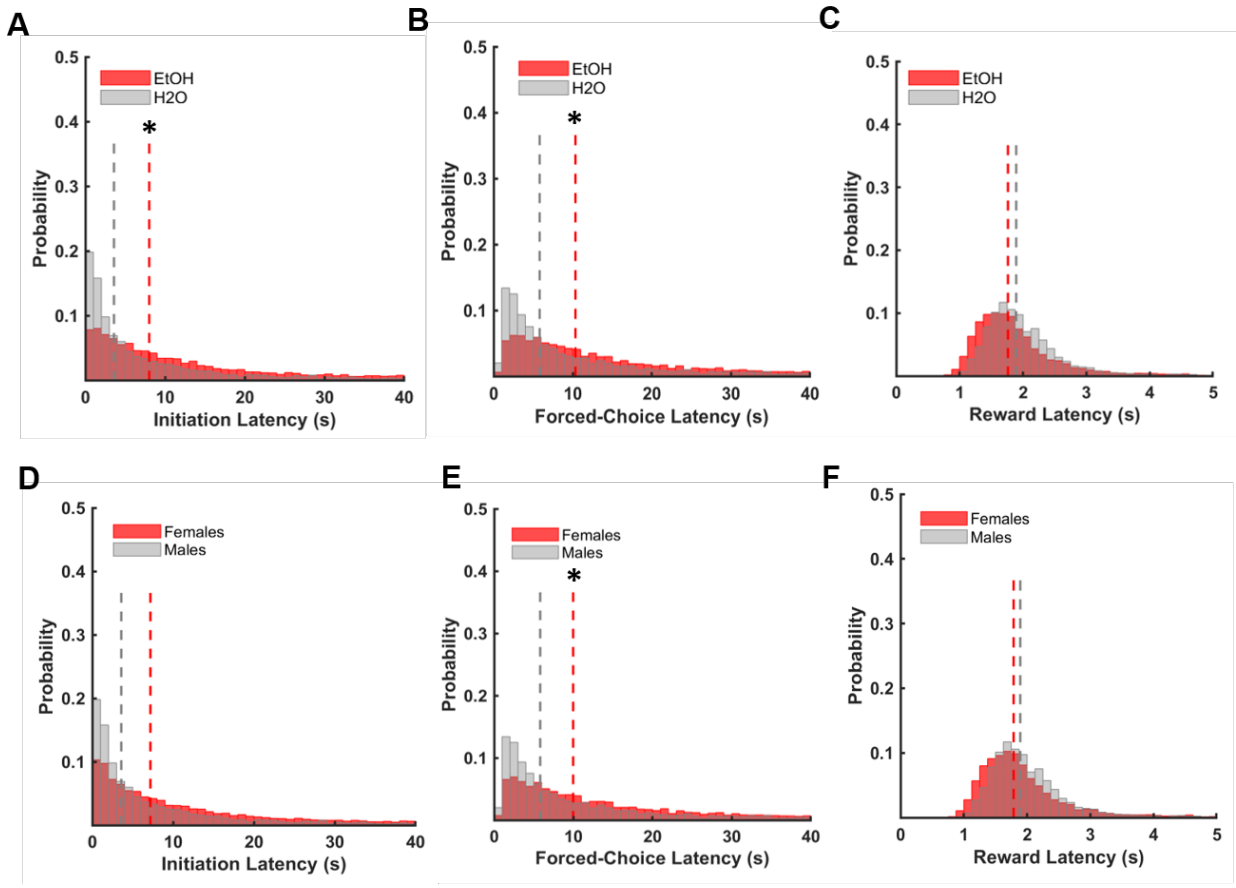

Figure 2-10. Drinking group and sex differences in latencies during operant pretraining (**A**) EtOH group exhibited longer initiation latencies than the H2O group. (**B**) EtOH group exhibited longer forced-choice latencies than the H2O group. (**C**) No group differences were found for reward latencies. (**D**) No sex differences were found for initiation latencies. (**E**) Females exhibited longer forced-choice latencies than males. **(F)** No sex differences were found for reward latencies. Dashed lines in latency histograms represent group medians. Bars indicate $\pm S.E.M.$ n=16 males, n=16 females, *p<0.05

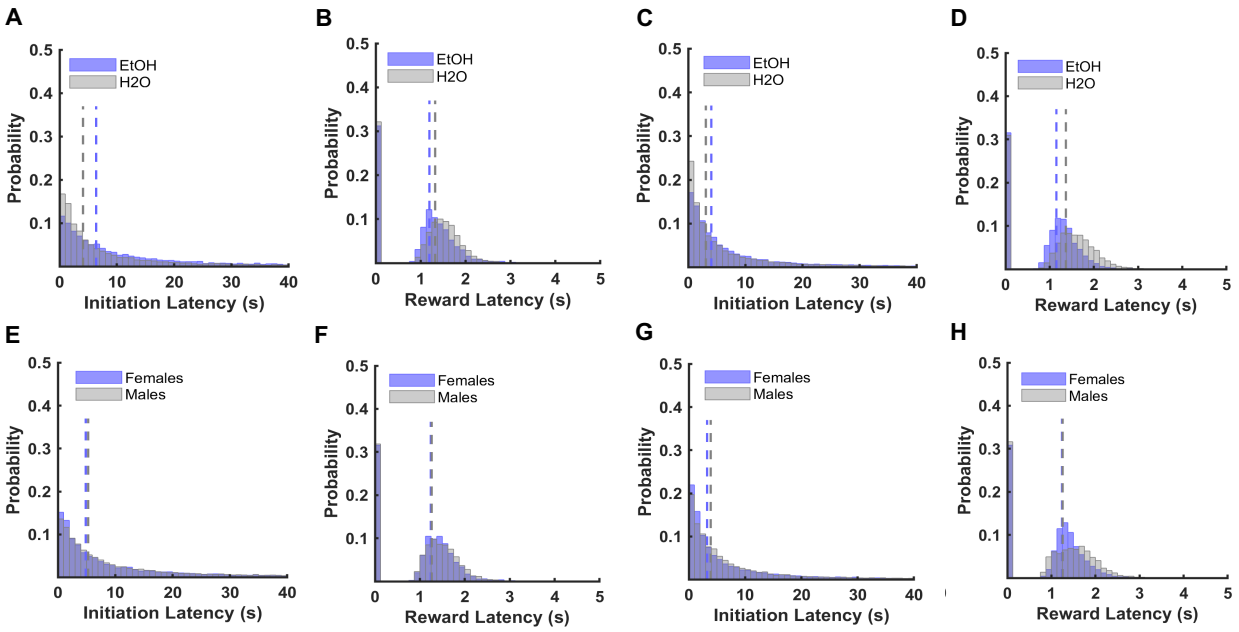

Figure 2-11. No drinking group or sex differences in latencies during early or late probabilistic discrimination learning. (**A**) No group differences were found for initiation latencies for the first 500 trials. (**B**) No group differences were found for reward latencies for the first 500 trials. (**C**) No group differences were found for initiation latencies for the last 500 trials. (**D**) No group differences were found for reward latencies for the last 500 trials. (**E**) No sex differences were found for initiation latencies for the first 500 trials. (**F**) No sex differences were found for reward latencies for the first 500 trials. (**G**) No sex differences were found for initiation latencies for the last 500 trials. (**H**) No sex differences were found for reward latencies for the last 500 trials. Dashed lines in histograms of latencies represent group medians. Bars indicate $\pm S.E.M.$ n=16 males, n=16 females.

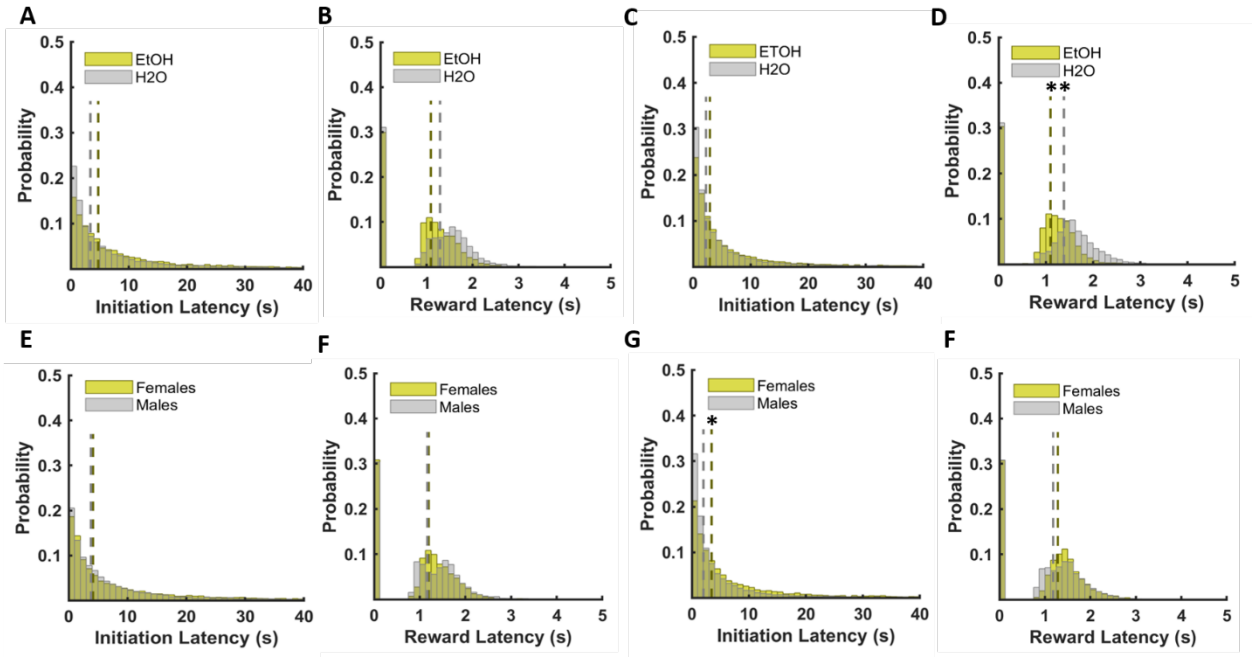

Figure 2-12. Drinking group and sex differences in latencies during early and late probabilistic reversal learning. (**A**) No group differences were found for initiation latencies for the first 500 trials. (**B**) No group differences were found for reward latencies for the first 500 trials. (**C**) No group differences were found for initiation latencies for the last 500 trials. (**D**) H2O group exhibited longer reward latencies compared to the EtOH group for the last 500 trials. (**E**) No sex differences were found for initiation latencies for the first 500 trials. (**F**) No sex differences were found for reward latencies for the first 500 trials. (**G**) Females exhibited longer initiation latencies compared to males for the last 500 trials. (**H**) No sex differences were found for reward latencies for the last 500 trials. Dashed lines in histograms of latencies represent group medians. Bars indicate $\pm\, S.E.M.$ n=16 males, n=14 females, *p ≤0.05, **p ≤0.01

Chapter 3: Unique features of stimulus-based probabilistic reversal learning

ABSTRACT

Reversal learning paradigms are widely-used assays of behavioral flexibility with their probabilistic versions being more amenable to studying integration of reward outcomes over time. Prior research suggests differences between initial and reversal learning, including higher learning rates, a greater need for inhibitory control, and more perseveration after reversals. However, it is not well-understood what aspects of stimulus-based reversal learning is unique to reversals, and whether and how observed differences depend on reward probability. Here, we used a visual probabilistic discrimination and reversal learning paradigm where male and female rats selected between a pair of stimuli associated with different reward probabilities. We compared accuracy, rewards collected, omissions, latencies, win-stay/lose-shift strategies, and indices of perseveration across two different reward probability schedules. We found that discrimination and reversal learning are behaviorally more unique than similar: fit of choice behavior using reinforcement learning models revealed a lower sensitivity to the difference in subjective reward values (greater exploration) and higher learning rates for the reversal phase. We also found latencies to choose the better option were greater in females than males but only for the reversal phase. Further, animals employed more win-stay strategies during early discrimination and increased perservation during early reversal learning. Interestingly, a consistent reward probability group difference emerged with a richer environment associated with longer reward collection latencies than a leaner environment. Future studies should systematically compare the neural correlates of fine-grained behavioral measures to reveal possible dissociations in how the circuitry is recruited in each phase.

INTRODUCTION

A critical feature of goal-directed, instrumental behavior is the ability to discriminate stimuli that predict reward from those that do not, and further, to flexibly update the response to those stimuli if predictions become inaccurate. Discrimination learning paradigms in rodents often involve pairing of an action (e.g., lever pressing, digging in a bowl, nosepoking stimuli on a touchscreen, or displacing an object) with an outcome (e.g., a desirable food reward). In typical paradigms, two or more stimuli are presented concurrently and the subject learns about the features of the stimuli that bring about reward and those that do not (e.g., nosepoking SA results in a better probability of reward than SB; pressing left lever yields better payout than right lever; scent A is more rewarded than scent B) (Alvarez & Eichenbaum, 2002; Dalton, Wang, Phillips, & Floresco, 2016; Eichenbaum, Fagan, & Cohen, 1986; Izquierdo et al., 2013; Schoenbaum, Chiba, & Gallagher, 2000; Schoenbaum, Nugent, Saddoris, & Setlow, 2002). With training, subjects become increasingly proficient at discrimination, in line with the associative rules imposed by the experimenter. The stimulus-reward rules can be deterministic (e.g., SA results in a sucrose pellet reward and SB does not) or probabilistic (e.g., SA results in a better probability of reward over SB), with deterministic and probabilistic schedules of reinforcement producing marked dissociations in the neural circuity recruited (Averbeck & Costa, 2017; Costa, Dal Monte, Lucas, Murray, & Averbeck, 2016).

In reversal learning paradigms (Izquierdo, Brigman, Radke, Rudebeck, & Holmes, 2017; Izquierdo & Jentsch, 2012), after either reaching a discrimination learning criterion for accuracy (Brushfield, Luu, Callahan, & Gilbert, 2008; Izquierdo et al., 2013; Stolyarova et al., 2019), a number of consecutive correct responses (Dalton, Phillips, & Floresco, 2014; Dalton et al., 2016), or a fixed block length of trials (Farashahi et al., 2017; Soltani & Izquierdo, 2019), the stimulus-

reward contingencies are reversed. At reversal, the trained response no longer results in a better probability of reward, though it usually remains the more frequently chosen option because of initial discrimination training. Indeed, usually reversals are acquired more slowly than the original discrimination, and younger subjects are quicker to learn than older ones (Brushfield et al., 2008; Schoenbaum, Setlow, Saddoris, & Gallagher, 2006). Perhaps partly due to this difference with original learning, reversal learning is considered unique in its requirement of flexibility, because it involves the subject inhibiting the prepotent response and, instead, responding to stimuli that were previously irrelevant. Other popular views are that discrimination and reversal learning phases occupy different task "spaces" (Wilson, Takahashi, Schoenbaum, & Niv, 2014) and that the phases differ in the likelihood that changes in contingencies will occur (Jang et al., 2015). Importantly, probabilistic learning and reversal paradigms, in particular, are more amenable to the application of reinforcement learning (RL) models that can estimate parameters for choice behavior based on the integration of previous rewarded and non-rewarded trials (Lee, Seo, & Jung, 2012). Despite various favored accounts, it is not well-understood how behaviorally unique reversal learning is compared to original (initial discrimination) learning. For example, latency measures can be used to dissociate attention, decision speed, and motivation via analyses of the time taken to initiate trials, choose a stimulus, and collect reward, respectively (Aguirre et al., 2020). Here, we explored if such detailed trial-by-trial measures of latencies and omissions differ across learning phases. Further, we probed if there are differences between reward probability schedules on these various measures by comparing the ability of separate cohorts of animals tested on different reward probability schedules to discriminate between the better and worse options, which were rewarded with a probability of 0.90 vs. 0.30, compared to 0.70 vs. 0.30. We also analyzed win-stay and lose-shift strategies, and perseveration and repetition metrics in each

learning phase. Finally, we investigated if RL models fit these learning phases differently. We studied this in both male and female animals.

We found a higher perseveration index and reduced use of win-stay strategies unique to early reversal compared to discrimination learning, which was expected. However, more surprisingly, we found consistent differences across discrimination and reversal learning phases to be limited to latencies to choose the better option (greater in females than males, only in the reversal) and to collect reward (greater in the higher reward probability group). These two metrics are proxies for decision speed and motivation, respectively. As for RL models, we found a lower sensitivity to the difference in subjective reward values and higher learning rates during reversal than initial discrimination. The sex differences, particularly in reversal learning, support previous findings using this paradigm. Interestingly, the only consistent probability group difference we observed across discrimination and reversal learning phases was in motivation (i.e. a richer environment was associated with longer reward collection latencies than a leaner environment). Collectively, our fine-grained analyses suggest that trial-by-trial behavioral measures of latencies and strategies may be particularly sensitive metrics to pair with neural correlate data in reversal learning. These measures may also be revealing in uncovering the unique substrates of flexible learning.

RESULTS

Comparisons of Sessions to Criterion, Omnibus measures

90-30 Reward Probability Control Cohorts. We first conducted statistical analyses to ensure the surgical control cohorts (c-cohort) [i.e., DREADD/VEH, eGFP (CNO+VEH)] that constituted the 90-30 reward probability group did not differ significantly on omnibus measures:

number of sessions to criterion (to 70% accuracy), probability of choosing the better option, and number of rewards collected during discrimination and reversal learning (i.e., first seven days of learning common to all rats). There was no effect of surgical control cohort (GLM: $\beta_{c\text{-}cohort}$ =4.17, t(5)=0.79, p=0.46) or sex (GLM: β_{sex} =8.17, t(5)=1.55 p=0.18) on sessions to reach criterion, probability correct (GLM: $\beta_{c\text{-}cohort}$ =0.004, t(54)=0.05, p=0.96; GLM: β_{sex} =0.06, t(54)=0.98, p=0.33), or number of rewards collected (GLM: $\beta_{c\text{-}cohort}$ =-6.39, t(54)=-0.27, p=0.78; GLM: β_{sex} =2.80, t(54)=0.09, p=0.93) for discrimination learning. Similarly, for reversal learning we did not find any effect of surgical control group (GLM: $\beta_{c\text{-}cohort}$ =-0.17, t(5)=-0.05, p=0.96) or sex (GLM: β_{sex} =-2.67, t(5)=-0.76, p=0.48) on sessions to reach criterion, probability correct (GLM: $\beta_{c\text{-}cohort}$ =0.02, t(53)=0.37, p=0.72; GLM: β_{sex} =0.10, t(53)=1.24, p=0.22), or number of rewards collected (GLM: $\beta_{s\text{-}group}$ =-18.52, t(53)=-1.19, p=0.24; GLM: β_{sex} =-7.81, t(53)=-0.40, p=0.69). Given the lack of differences between surgical control cohorts within the 90-30 reward probability group, we collapsed data across these cohorts for further analyses.

Number of sessions. In total, the 70-30 probability group performed a total of 847 sessions, 417 during the discrimination phase and 430 during reversal (an average of 26.1 $\pm$ 3.1 during the discrimination phase and 26.9 $\pm$ 1.71 during the reversal phase). The 90-30 probability group performed 270 sessions, 118 during discrimination and 152 during reversal (an average of 13.1 $\pm$ 2.6 during discrimination and 16.9 $\pm$ 1.65 reversal), **Figures 3-2B** and **3-2D**. For discrimination learning, we found no effect of group (p=0.24), sex (p=0.21), or significant group*sex interaction (p=0.49) on sessions to reach criterion. The 90-30 reward probability group required an average of 11.56 $\pm$ 2.56 (M $\pm$ SEM) sessions while the 70-30 reward probability group required an average of 21.13 $\pm$ 3.26 sessions to reach a 70% criterion; females required an average of 11.92 $\pm$ 2.47 sessions, while males required an average of 23.92 $\pm$ 3.59 sessions. For reversal learning,

there was a significant effect of group (GLM: β_{group} =8.28, t(21)=2.63, p=0.02), with the 90-30 reward probability group requiring fewer sessions to reach criterion on average (17.44 $\pm$ 1.68) than the 70-30 reward probability (23.50 $\pm$ 1.95). However, there was no effect of sex (p=0.20), with males (17.67 $\pm$ 1.97) and females (24.69 $\pm$ 1.80) requiring a comparable number of sessions to reach criterion for reversal learning, and no significant group*sex interaction (p=0.41). As differing rates of acquisition during discrimination learning could be attributed to differences in performance in the reversal learning phase, discrimination sessions to criterion was included as a covariate in reversal learning analyses (whenever a phase interaction justified analysis of each phase separately), specifically on the main behavioral outcome measures for which those interactions were found. As a preview, the pattern of results were largely consistent with those obtained without the covariate in the model.

Comparisons of Accuracy and Rewards Collected

Comparisons between early Discrimination & Reversal Learning. We fitted GLMs that combined analysis of the first 7 days of learning across both phases (discrimination and reversal). For *probability of choosing the better option* (**Appendix B Table 1**)*,* we found that all rats exhibited learning by demonstrating an increase in choosing the better option across days (**Figure 3-3**). Overall animals chose the better option more in the discrimination phase than the reversal phase. For the *number of rewards* (**Appendix B Table 2**), all rats increased their number of rewards collected across days for both learning phases. There were no difference between reward probability group , or learning phase, nor were there any factor interactions on probability correct or number of rewards. As there were no significant phase interactions, we were not justified to analyze the learning phases separately for these measures.

Summary. The omnibus comparison across phases of early discrimination and reversal learning revealed that all animals demonstrated an increase in accuracy (i.e. probability of choosing the better option) and an increase in the number of rewards collected across days, regardless of reward probability group or sex. However, animals chose the better option more often in the discrimination phase, compared to reversal as expected.

Comparisons of Omissions and Latencies

Comparisons between early Discrimination & Reversal Learning. Similar to above, we fitted GLMs that combined analysis of the first 7 days of learning across both phases (discrimination and reversal), see **Appendix B Table 3**. For total number of *initiation omissions*, there was a significant phase*group*sex interaction, with post-hoc comparisons revealing 70-30 males exhibited more initiation omissions in the discrimination phase than the reversal phase ($p=0.01$). There was also a significant group*sex interaction on this measure, but post-hoc tests were not significant after accounting for the number of comparisons. There was no effect of phase, reward probability group, or sex, and no significant phase*group, or phase*sex interactions on this measure. For total number of *choice omissions*, there was no effect of phase, reward probability group, or sex, and no significant phase*group, phase*sex, group*sex, or phase*group*sex interactions.

Next we analyzed median initiation and better choice latencies (**Appendix B Table 4**), as well as worse choice and reward collection latencies (**Appendix B Table 5)**. For *initiation latencies*, we found a marginal group*sex interaction, but no effect phase, or reward probability group, and no significant phase*group, phase*sex, group*sex, or phase*group*sex interactions. For *better choice latencies*, there was a significant phase*sex interaction, with males exhibiting

longer choice latencies for the better option in the discrimination phase compared to the reversal phase (p=0.01), but no effect of phase, reward probability group, or sex, and no significant phase*group, group*sex, or phase*group*sex interactions. For *worse choice latencies*, there was a significant effect of phase, with animals in the reversal phase exhibiting longer latencies than in the discrimination phase, but no effect of reward probability group, sex, or significant phase*group, phase*sex, group*sex, phase*group*sex interactions. For *reward collection latencies*, we found a significant effect of reward probability group, with the 90-30 group taking longer to collect reward than the 70-30 group, but no effect of phase or sex. There was a significant phase*sex interaction, with females exhibiting longer latencies than males in the reversal phase (p=0.01), and a phase*group*sex interaction, with 90-30 males in the discrimination phase than 90-30 males in the reversal phase, but no significant phase*group or group*sex interaction.

Comparisons within Discrimination Learning. Because we found phase interactions on initiation omissions, better choice latencies, and reward collection latencies, we used GLMs to analyze these variables for the discrimination phase separately (**Appendix B Table 6**). For total number of *initiation omissions*, there was no significant effects of reward probability group, sex, or group*sex interaction (**Figure 3-4A**). For *better choice latencies*, there was also no effect of reward probability group or sex, or group*sex interaction, despite a phase by sex interaction found during early learning (**Figures 3-4B, 3-6A, and 3-6B**). However, for *reward collection latencies*, there was a significant effect of reward probability group with the 90-30 group taking longer to collect reward than the 70-30 group, but no effect of sex and no group*sex interaction (**Figures 3-4C, 3-6C, and 3-6D**).

Comparisons within Reversal Learning. As above, because we found phase interactions on initiation omissions, better choice latencies, and reward collection latencies, we used GLMs to

analyze these variables for the reversal phase separately. We ran two models for the reversal learning phase: an unadjusted model which included only the main factors (i.e. *group* and *sex*), and an adjusted model with the number of discrimination sessions to criterion added as a covariate (**Appendix B Tables 7 and 8**). For total number of *initiation omissions*, we found a marginal effect of reward probability group (p=0.07) in the adjusted model, but did not find a significant sex, or group*sex interaction with either model (**Figure 3-4D)**. For *better choice latencies*, there was an effect of sex in both the unadjusted and adjusted model, with females exhibiting longer choice latencies than males when choosing the better option. No significant group or group*sex interactions were found for either model (**Figures 3-4E, 3-6E, and 3-6F**). For *reward collection latencies*, there was a significant effect of reward probability group, with the 90-30 group taking longer to collect reward than the 70-30 group, in both the unadjusted and adjusted model, but no significant sex or group*sex interaction was observed (**Figures 3-4F, 3-6G and 3-6H**).

Summary. After controlling for the number of discrimination sessions to criterion in reversal learning, we observed the same pattern of effects as that obtained with the original model: female animals exhibited longer choice latencies for the better option than males (this pattern was not observed for the discrimination phase). Additionally, we found longer reward collection latencies in animals learning the 90-30 reward probabilities, compared to animals in the 70-30 group. This effect was observed in both the discrimination and reversal phases. Finally, though we initially found a phase interaction for initiation omissions in early learning, follow-up analyses yielded only a marginal effect of group, but no sex, or group by sex interactions when the phases were analyzed separately. In sum, we determined that latencies, and not omissions, are among the more sensitive measures of performance; certainly beyond omnibus measures of accuracy and cached rewards that are typically reported in the literature.

Comparisons of Response to Reward Feedback

Comparisons between early Discrimination & Reversal Learning. We analyzed win-stay and lose-shift strategies during early discrimination and reversal learning (**Appendix B Table 9**). For win-stay, we found a marginally significant effect of phase on employing the win-stay strategy, with animals using this strategy more in the discrimination phase than in the reversal phase, a significant effect of reward probability group, with animals in the 90-30 reward probability group using this strategy more than those in the 70-30 reward probability group, but no effect of sex. There was also a significant phase*group interaction, with animals in the 90-30 reward probability group using this strategy more in the discrimination phase than the reversal phase ($p=0.001$), but no significant phase*sex, group*sex, or phase*group*sex interactions. For lose-shift, there was a significant effect of phase, with animals employing this strategy more in the discrimination learning phase than the reversal phase, a significant effect of reward probability group, with the 90-30 group using the lose-shift strategy more than the 70-30 group. Finally, there was also an effect of sex, with males employing this strategy more than females, but no significant phase*group, phase*sex, group*sex, phase*group*sex interactions.

Comparisons within Discrimination Learning. Because we found phase interactions on win-stay, we were justified to analyze potential differences in win-stay strategies on stimulus responses employed by each reward probability group and by sex, in the discrimination phase separately (**Appendix B Table 10)**. For win-stay, there was an effect of reward probability group, but no effect of sex, and no group*sex interaction (**Figure 3-5A**). When we considered win-stay *on the better option* (win-stay|better), we found a marginally significant effect of reward probability group, with the 90-30 group employing this strategy more, but no effect of sex or significant group*sex interaction (**Figure 3-5B**). When we analyzed win-stay on the worse option

(win-stay|worse), we found no significant effect of reward probability group, sex, or group*sex interaction (**Figure 3-5C**).

Comparisons within Reversal Learning. Per the phase interactions, we were permitted to analyze win-stay (**Appendix B Table 11**) strategies on stimulus responses employed during the reversal phase, separately. We found males were less likely to employ a win-stay strategy (**Figure 3-5D)**, but no effect of reward probability group, or group*sex interaction. We found no significant group or sex differences, and no group*sex interaction for win-stay|better (**Figure 3-5E**). We did, however, find a significant effect of sex on win-stay|worse, with males less likely to employ the strategy than females, but no reward probability group effect or group*sex interaction (**Figure 3-5F**).

Summary. Win-stay and win-stay|better strategies were employed more often during early discrimination learning compared to early reversal learning. Further, the 90-30 reward probability group used both win-stay and win-stay|better strategies more in discrimination learning than the 70-30 group. Lastly, female animals were more likely to apply a win-stay|worse strategy in the reversal phase compared to males.

Comparisons of Repetition measures

Refer to Appendix A.

Comparisons of Estimated Parameters based on Fit of Choice Behavior with RL models

Refer to Appendix A.

DISCUSSION

Here, we used a stimulus-based probabilistic discrimination and reversal paradigm with different probabilities of reward (i.e., 90/30, 70/30) to test learning and performance on several

measures (i.e., probability of choosing the better option, initiation and choice omissions and latencies, perseveration and repetition measures, and win-stay/lose-shift strategies). We also examined whether fit of choice data using RL models would reveal differences between the two learning phases, and tested for potential differences in learning rates (single, separate) and explore/exploit behavior. We found higher learning rates in the reversal than discrimination phase, which mirror recent reports (described in more detail below). Animals also exhibited decreased sensitivity to the difference in subjective reward values of the two options during the reversal learning phase compared to discrimination, indicative of greater exploration. Further, we found increased perseveration in early reversal compared to early discrimination learning. Finally, we found that differences in reward collection latencies depended on the richness of the environment (90% vs 70% reward). Notably, the only pronounced sex difference was in longer latencies to choose the better option during reversal learning (females taking longer than males), which is generally consistent with the pattern of effects of a recent study by our group (Aguirre et al., 2020). We elaborate on these findings in the context of the broader literature below.

Learning rates

We found that the learning rate was higher in reversal than discrimination learning. The general trend of increased learning rates following reversal and decreased sensitivity to the difference in subjective reward values is consistent with existing literature (Costa et al., 2015; Massi et al., 2018). The increased learning rate suggests a more rapid integration of reward feedback, while the lower sensitivity to difference in reward values indicates rats choose the higher-valued option less consistently, corresponding to greater exploration. These changes may reflect response to increased environmental volatility following the reversal in terms of both faster learning and more exploration. Bathellier et al. (2013) accounted for a similar change in learning

rate in mice by assuming slower initial learning during discrimination due to weaker initial synaptic weights, and much faster learning during reversal due to the same synapses being activated, but at a state where synaptic weights are stronger than they were in discrimination. Future modeling studies are needed to explain how a single reversal can induce both higher learning rates and decreased sensitivity to subjective values.

Latencies to collect reward and choose the better option

Probability group differences across both learning phases were found for reward collection latencies, with the 90-30 group exhibiting longer reward collection times than the 70-30 group, suggesting the 90-30 group may experience attenuated motivation by comparison. Latency to collect reward is commonly used as a measure of motivation, whereas stimulus response latency (i.e. latency to choose the better option) is often used as a measure of processing or decision-making speed in the Five-Choice Serial Reaction Time Task (5CSRTT) (Amitai & Markou, 2010; Asinof & Paine, 2014; Bari et al., 2008; Bushnell & Strupp, 2009; Chudasama et al., 2003; Remmelink et al., 2017; Robbins, 2002; Robinson et al., 2009). One interpretation of this finding is that animals may have been more motivated to confirm whether they had received a reward under higher uncertainty (i.e., when the probability for the better option was lower), and conversely, were more confident in their decision with a higher reward probability associated with choosing the better option. Another related explanation is that animals exert more effort and display more vigor when in leaner and more uncertain reward environments (Amsel, 1967; McNamara et al., 2013), which in our experiments was directly tied to the differences in reward probabilities associated with the better option. Because there was greater reward uncertainty associated with the better option for the 70-30 group, rats may have exerted more effort to retrieve

rewards compared to the 90-30 group. Interestingly, although animals generally increased their choice of the better option across days during discrimination learning, this was actually not as much the case for reversal learning where this measure over time was relatively flat. Despite the reward collection latency differences, we found no difference in the number of cached rewards between the probability groups, so these differences are likely not due to satiety *per se*, but instead, an effect of an experience with overall reward rate over time.

A final consideration related to our consistent probability group difference is a recent report (Song & Lee, 2020) providing evidence for differential neural recruitment in environments that require different "resource allocations." Agents (or in our case, rats) learn over time to assign resources to certain stimuli, and make adjustments as stimuli gain more reward-predictive value. The leaner reward schedule (70-30) may actually promote a more adjustable, flexible resource allocation than the more profitable schedule (90-30). Empirically testing how midbrain dopamine interacts with cortical structures to support flexible resource allocation is an interesting line of future inquiry. To probe this, different reward probability schedules could be compared.

Perseveration and Win-Stay Strategies

All animals exhibited more perseveration during early reversal learning than early discrimination learning. This finding supports those by Verharen et al. (2020) in which they simulated data of thousands of probabilistic reversal learning sessions using a Q-learning model consisting of separate learning rates for learning from positive (i.e. rewarded) and negative (i.e. nonrewarded) feedback, a beta parameter, and a stickiness parameter indicative of perseveration. They found that a greater number of reversals occurred when the stickiness parameter value was high (i.e., greater perseveration), but only when both learning rates were also high (Verharen et

al., 2020). Interestingly, we found reward probability group differences for perseveration, with the 90-30 group exhibiting greater perseveration than the 70-30 group. This finding implies that more consistent reward feedback (i.e. higher reward probability associated with the better option) in the 90-30 group promoted perseverative responding in early learning. Furthermore, males demonstrated greater exploratory behavior (i.e. lower perseveration) during discrimination learning, in line with previous research showing greater impulsivity in male rats (Lukkes et al., 2016; Palanza et al., 2001).

Greater perseveration during early reversal is consistent with the long-standing idea that reversal learning is a measure of inhibitory control, such that the inability to disengage from previously rewarding behavior after a change in contingency may be reflective of compulsive and even impulsive response tendencies, commonly associated with drug dependence (Izquierdo & Jentsch, 2012). Indeed, there is evidence that inflexible responding in reversal learning may be genetically related to impulsivity (Crews & Boettiger, 2009; Fineberg et al., 2010; Franken et al., 2008; Groman et al., 2009; Groman & Jentsch, 2013; Jentsch et al., 2014). However, greater perseveration can also be explained by slower updating of model-free learning in which animals are not benefiting as much from trial-by-trial feedback, or (just as likely) failing to detect task transitions, especially in early reversal (Izquierdo et al., 2017). It is particularly interesting that we show here *stimulus* perseveration, aside from spatial- or response-based perseveration, which we control for here by pseudorandomly presenting the stimuli on the left- vs. right- sides of the screen.

A related observation we report here is that animals more often used reward-dependent choice strategies (i.e., win-stay and win-stay|better) in the early discrimination phase compared to the reversal phase, and adopted an opposing pattern after reversal (i.e., win-stay|worse as more prevalent). Importantly, as above, this strategy is stimulus-dependent and not location-dependent,

which is more often probed in lever-based (left vs. right) tasks. Indeed, animals exhibited less consistent response to reward feedback during the reversal compared to the discrimination phase, indicative of noisier behavior or equivalently more exploration.

Sex differences

Females exhibited longer latencies to choose the better option than males, and this was only observed during the reversal learning phase. Measures of decision speed in rodents vary, but can include latencies to nosepoke (i.e., time to make a response before the end of the trial), as well as percentage of correct responses (i.e., accuracy), usually on the 5CSRTT as described above. Our data supports the interpretation that females exhibit greater response demands (and consequently slower performance) than males in the reversal phase. This adds to a growing literature on sex differences in reversal learning (Aarde et al., 2020; Bissonette et al., 2012; Branch et al., 2020; LaClair & Lacreuse, 2016).

Conclusion

The present results suggest that certain measures of decision speed (i.e., choice latencies) and motivation (i.e. reward collection latencies) should be used as more than auxiliary measures to study reversal learning. Indeed, latencies, and not omissions, are more sensitive measures than omnibus measures of accuracy and cached rewards that are typically reported in the literature. Some of the measures we studied here are likely correlated with others and it would be interesting in follow-up experiments to pinpoint the most predictive factors in discriminating the two types of learning using a much larger dataset. Future studies should probe the neural correlates of these

fine-grained behavioral measures, as these have been under-utilized and may reveal marked dissociations in how the circuitry is recruited in each phase.

METHODS

Subjects

Subjects were N=25 adult male (n=13) and female (n=12) Long-Evans rats (Charles River Laboratories) aged > post-natal-day (PND) 60 at the start of testing. Rats arrived to the vivarium between PND 40-60. The rats included in this report served as controls for two different experiments: one was a cohort of water (H_2O)-only drinking (n=8 female and n=8 male) rats that served as controls in an ethanol study (see *2-Bottle Choice*) and the other was a cohort of rats (n=5 female and n=4 male) that experienced surgical procedures (see *Surgery*), serving as controls for a study targeting the orbitofrontal cortex (OFC) with DREADDs. Importantly, all rats were the same age (PND 140-155) when pretraining commenced, and further, all rats were part of experiments that ran in parallel, minimizing differences between cohorts.

Before any treatment, all rats underwent a 3-day acclimation period during which they were pair-housed and given food and water *ad libitum*, and remained in cages with no experimenter interference. Following this 3-day acclimation period, animals were handled for 10 min per animal for 5 consecutive days. During the handling period, the animals were also provided food and water *ad libitum*. After the handling period, animals were individually-housed under standard housing conditions (room temperature 22–24° C) with a standard 12 h light/dark cycle (lights on at 6am). Following either 2-bottle choice or surgery, rats were tested on probabilistic discrimination and reversal learning, as below. All procedures were conducted in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National

Institutes of Health and the Chancellor's Animal Research Committee at the University of California, Los Angeles.

2-Bottle Choice

Home cages were modified to allow for the placement of two bottles for drinking, whereas standard housing allows for only one bottle. Rats (n=16; 8 male and 8 female) included in this analysis were singly-housed and given access to 2 H_2O bottles simultaneously (no ethanol) for 10 weeks, the same duration that experimental animals were provided the choice of ethanol vs. H_2O. Weight of bottles was measured three times per week to measure consumption amounts, compared to a control cage placed on the same rack to account for leakage. Rats were not monitored for weight during this time.

Surgery

Viral Constructs

Rats (n=9; 4 male and 5 female) included in the present comparison were singly-housed and allowed to express DREADDs or eGFP in OFC for 6 weeks, the same duration experimental animals that were treated with clozapine-N-oxide (CNO) were allowed to express DREADDs. An adeno-associated virus AAV8 driving the hM4Di-mCherry sequence under the CaMKIIa promoter was used to express DREADDs on OFC neurons (AAV8-CaMKIIa-hM4D(Gi)-mCherry, packaged by Addgene) (Addgene, viral prep #50477-AAV8). A virus lacking the hM4Di DREADD gene and only containing the green fluorescent tag eGFP (AAV8-CaMKIIa-EGFP, packaged by Addgene) was also infused into OFC in separate cohorts of animals as a null virus control. Altogether the animals included in these sets of analyses served as control cohorts for a larger experiment in which they were given subcutaneous injections of either CNO or a saline

vehicle (VEH) prior to reversal learning. The animals included here were 5 animals prepared with hM4Di DREADDs in OFC who received VEH, 2 animals prepared with eGFP in OFC who received VEH, and 2 animals prepared with eGFP in OFC, who received CNO during reversal learning. Importantly, although these animals received virus in OFC, the DREADDs were not activated. Further, we provide analyses to show these cohorts did not differ, and could be combined.

Surgical Procedure

Infusion of AAV virus containing DREADD or eGFP (n=9) in OFC was performed using aseptic stereotaxic techniques under isoflurane gas (1-5% in O_2) anesthesia prior to any behavioral testing experience. Before surgeries were completed, all animals were administered 5mg/kg s.c. carprofen (NADA #141–199, Pfizer, Inc., Drug Labeler Code: 000069) and 1cc saline. After being placed in the stereotaxic apparatus (David Kopf; model 306041), the scalp was incised and retracted. The skull was leveled to ensure that bregma and lambda were in the same horizontal plane. Small burr holes were drilled in the skull above the infusion target. Virus was bilaterally infused at a rate of 0.02 µl per minute for a total volume of 0.2 µl per hemisphere into OFC (AP = +3.7; ML= ±2.0; DV = −4.6, relative to bregma). After each infusion, 10 min elapsed before the syringe was pulled up.

Food Restriction

Five days prior to any behavioral testing, rats were placed on food restriction with females on average maintained 12-14 grams/ day and males given 16-18 grams/ day of chow. Food restriction level remained unchanged throughout behavioral testing, provided animals completed

testing sessions. Water remained freely available in the home cage. Animals were weighed every other day and monitored closely to not fall below 85% of their maximum, free-feeding weight.

Learning

Pretraining. Behavioral testing was conducted in operant conditioning chambers outfitted with an LCD touchscreen opposing the sugar pellet dispenser. All chamber equipment was controlled by customized ABET II TOUCH software.

The pretraining protocol, adapted from established procedures (Stolyarova & Izquierdo, 2017), consisted of a series of phases: Habituation, Initiation Touch to Center Training (ITCT), Immediate Reward Training (IMT), designed to train rats to nosepoke, initiate a trial, and select a stimulus to obtain a reward (i.e. sucrose pellet). During habituation, rats were required to eat five pellets out of the pellet dispenser inside the chambers within 15 min before exposure to any stimuli on the touchscreen. ITCT began with the display of white graphic stimuli on the black background of the touchscreen. During this stage, a trial could be terminated for one of two reasons: if a rat touched the displayed image and received a reward, or if the image display time (40 s) ended, after which the stimulus disappeared, a black background was displayed, and a 10 s inter-trial interval (ITI) ensued. If the rat did not touch within 40 s this was scored as an *initiation omission*. IMT began in the same way as ITCT, but the disappearance of the white graphic stimulus was now paired with the onset of a target image immediately to the left or right of the stimulus (i.e. forced-choice). During this stage, a trial could be terminated for one of three reasons. First, if a rat touched the center display (i.e. white graphic stimulus) and touched the image displayed on either side, after which there was a dispensation of one sucrose pellet and illumination of the tray-light. Second, if the rat failed to touch the center white graphic stimulus after the display time ended (40

s), after which the stimulus disappeared, a black background was displayed, and a 10 s ITI ensued, scored as an *initiation omission*. Third, if the image display time (60 s) ended, after which the stimulus disappeared, a black background was displayed, and a 10 s ITI ensued, scored as a *choice omission*. Rats could also fail to respond to the center stimulus within 40 s during this phase (i.e. initiation omission, as in the previous phase). For habituation pretraining, the criterion for advancement was collection of all 5 sucrose pellets. For ITCT, the criterion to the next stage was set to 60 rewards consumed in 45 min. The criterion for IMT was set to 60 rewards consumed in 45 min across two consecutive days.

Probabilistic Discrimination Learning. After completion of all pretraining schedules, rats were advanced to the discrimination (initial) phase of the PRL task, in which they would initiate a trial by touching the white graphic stimulus in the center screen (displayed for 40 s), and choose between two visual stimuli presented on the left and right side of the screen (displayed for 60 s) counterbalanced between trials, assigned as the better or worse options, with a reward (i.e. sucrose pellet) probability of either $p_R(B)$= 0.90 or 0.70 (i.e. better option) or $p_R(W)$=0.30 (i.e. worse option). If a trial was not initiated within 40 s, it was scored as an initiation omission. If a stimulus was not chosen, it was scored as a choice omission, and a 10 s ITI ensued. If a trial was not rewarded, a 5 s time-out would follow, subsequently followed by a 10 s ITI. Finally, if a trial was rewarded, a 10 s ITI would follow after the reward was collected (**Figure 3-1**). The criterion was set to 60 or more rewards consumed and selection of the better option in 70% of the trials or higher during a 60 min session across two consecutive days. After reaching the criterion for the discrimination phase, the rats advanced to the reversal phase beginning on the next session. Notably, one animal in the 90-30 group and five animals in the 70-30 group did not meet discrimination criterion and were forced-reversed after 25+ days.

Probabilistic Reversal Learning. After the discrimination phase, the rats advanced to the reversal phase during which rats were required to remap stimulus-reward contingencies and adapt to reversals in the reward probabilities. The stimuli associated with the p_R (B)=0.90 or 0.70 probability (i.e. better option), would now be associated with a p_R(W)=0.30 probability of being rewarded (i.e. worse option). Consistent with prior literature showing freely-behaving rodents exhibit slow learning on probabilistic reversals with visual stimuli (Aguirre et al., 2020), most animals from either cohort did not meet a 70% criterion before the termination of the study, so we limited our analyses to the first seven sessions of discrimination and reveral phases for all animals.

Data Analyses

MATLAB (MathWorks, Natick, Massachusetts; Version R2019b) was used for all statistical analyses and figure preparation. Data were analyzed with a series of mixed-effects General Linear Models (GLM); omnibus analyses across discrimination and reversal phases in early learning (operationally defined as the first seven sessions), and then individual analyses within each phase separately if justified by a phase interaction. We and others have analyzed early learning in previous work, as it may be particularly informative to revealing sensitivity to reward feedback and perseveration (Izquierdo et al., 2010; Jones & Mishkin, 1972; Stolyarova et al., 2014; Stolyarova et al., 2019), measured using touchscreen response methods (Izquierdo et al., 2006). In the individual analyses for the reversal learning phase, we first ran an unadjusted model which only included the main factors (i.e. *day, group,* and *sex*) and their interactions for our main behavioral outcome measures (i.e. *probability of choosing the better option, number of rewards, omissions*, and *latencies*) for which a phase interaction was obtained. This was followed by an adjusted model, which included discrimination sessions to criterion as a covariate. The adjusted

model was generated to ensure that differences between discrimination and reversal measures were not due to training differences between groups or within individual animals.

Learning data were analyzed with GLM (*fitglme* function; Statistics and Machine Learning Toolbox; MathWorks, Natick, Massachusetts; Version R2017a), with learning phase (discrimination vs reversal), probability group (90-30 vs 70-30), and sex (male vs female) as fixed factors, and individual rat as random factor. All Bonferroni post-hoc tests were corrected for number of comparisons. Statistical significance was noted when p-values were less than 0.05, and p-values between 0.05 and 0.07 were reported as a trend, or marginally significant. Major dependent variables include: probability correct, number of rewards (sucrose pellets earned), total initiation omissions (failure to initiate a trial), total choice omissions (failure to select a stimulus), and median latencies (to initiate a trial, to nosepoke the correct stimulus, to nosepoke the incorrect stimulus, and to collect reward). The latter we refer to as initiation-, correct-, incorrect-, and reward latencies, respectively.

Each trial was classified as a *win* if an animal received a sucrose pellet, and as a *loss* if no reward was delivered. Decisions were classified as better if the animal chose the more rewarding stimulus (stimulus with the larger probability of reward) and *worse* if it chose the less rewarding stimulus. We classified decisions as *Stays* when a rat chose the same stimulus on the subsequent trial and as *Shifts* when it switched to the other alternative. From these first-order measures we were able to construct win-stay, the probability of choosing the same stimulus on the following trial after being rewarded, and lose-shift, the probability of choosing the alternative stimulus after not receiving a reward. These were further parsed into better or worse win-stay or lose-shift, depending on whether the win-stay/lose-shift followed selection of the better option or the worse option. Because we were primarily interested in the differences between the initial phases of

discrimination and reversal learning, only the first seven sessions (i.e., early phase learning) for each animal were included in our analysis on response to reward feedback.

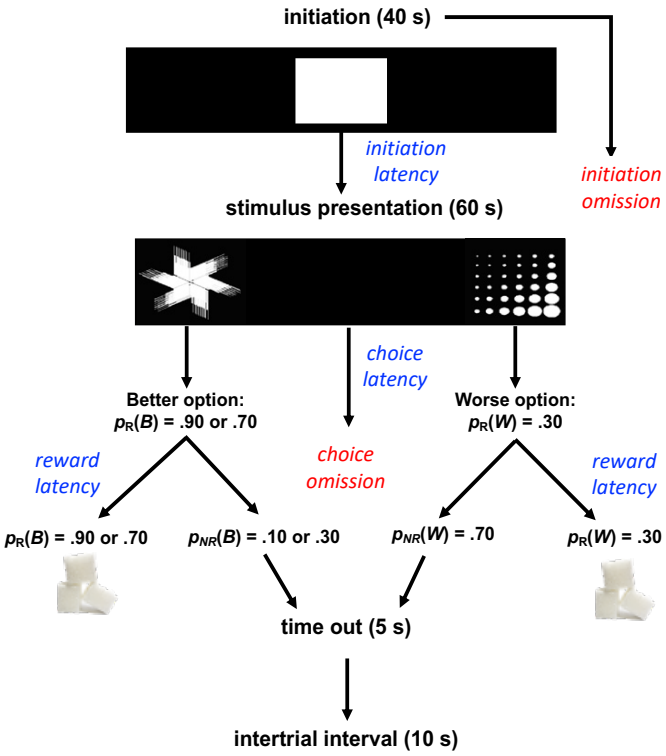

Figure 3-1. Task design. Schematic of probabilistic learning task. Rats initiated a trial by nosepoking the center stimulus (displayed for 40 s) and then selected between two visual stimuli pseudorandomly that were presented on either the left and right side of the screen (displayed for 60 s), assigned as the better (B) and worse (W) options. Correct nosepokes were rewarded with a sucrose pellet with probability $p_R(B)$=0.90 or 0.70 versus $p_R(W)$=0.30. If a trial was not rewarded [$p_{NR}(B)$ or $p_{NR}(W)$], a 5 s time-out would commence. If a stimulus was not chosen, it was considered a choice omission and a 10 s ITI would commence. Rats could also fail to initiate a trial, in which case, it was scored as an initiation omission. If a trial was rewarded, a 10 s ITI would follow reward collection. Other prominent measures collected on a trial-by-trial basis were trial initiation latency (time to nosepoke the center white square), choice latency (time to select between the two stimuli), and reward latency (time to collect reward in the pellet magazine).

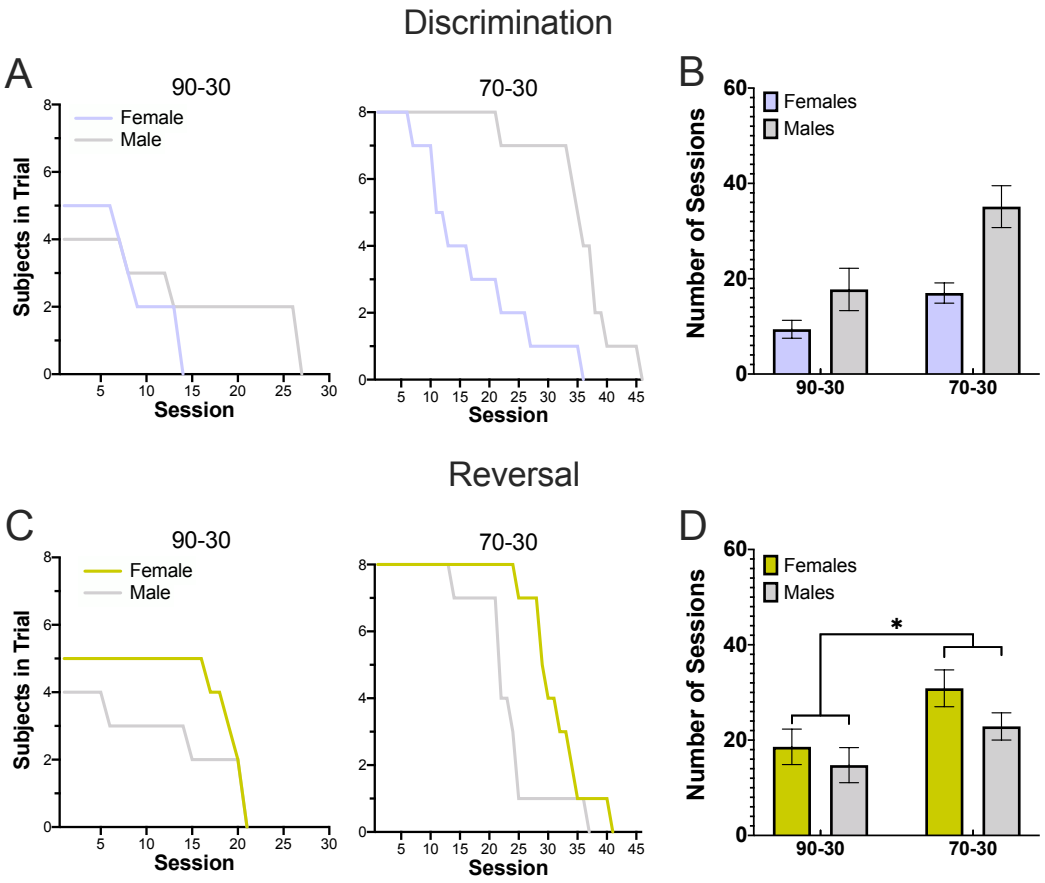

Figure 3-2. Greater number of completed sessions for the 70-30 reward probability group in both discrimination and reversal learning. (A-B) Plotted are the number of subjects per session (A) and the number of sessions to criterion (B) during the discrimination (pre-reversal) phase. The 70-30 reward probability group completes significantly more sessions during discrimination than the 90-30 group. **(C-D)** The same as A-B but during for reversal learning. The 70-30 reward probability group again completes significantly more sessions than the 90-30 group. Bars indicate $\pm$ *SEM* *p $\leq$0.05

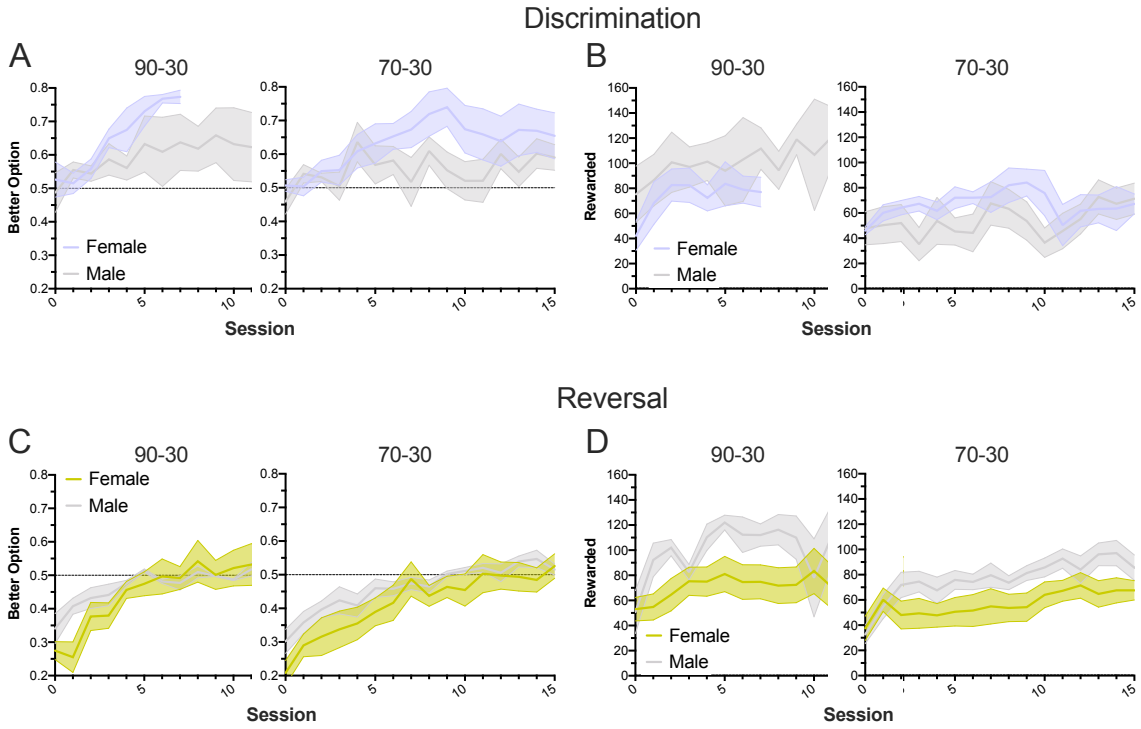

Figure 3-3. Both reward probability groups and both sexes increase their collected rewards over time but animals choose the better option more often in the discrimination phase. (A-B) Proportion of better option selections (A) and number of rewards in a session (B) in the discrimination (pre-reversal) phase, showing the first 10 and 15 trials. Both groups increase selection of the better option and receive more rewards per session over time, with no significant differences between reward probability groups or sex. **(C-D)** Same as A-B, but in the reversal phase. Again, animals in both reward probability groups improve accuracy and collected rewards over time, with no differences by group or sex. Notably, there was significant phase difference on choice of the better option, with the discrimination > reversal phase.

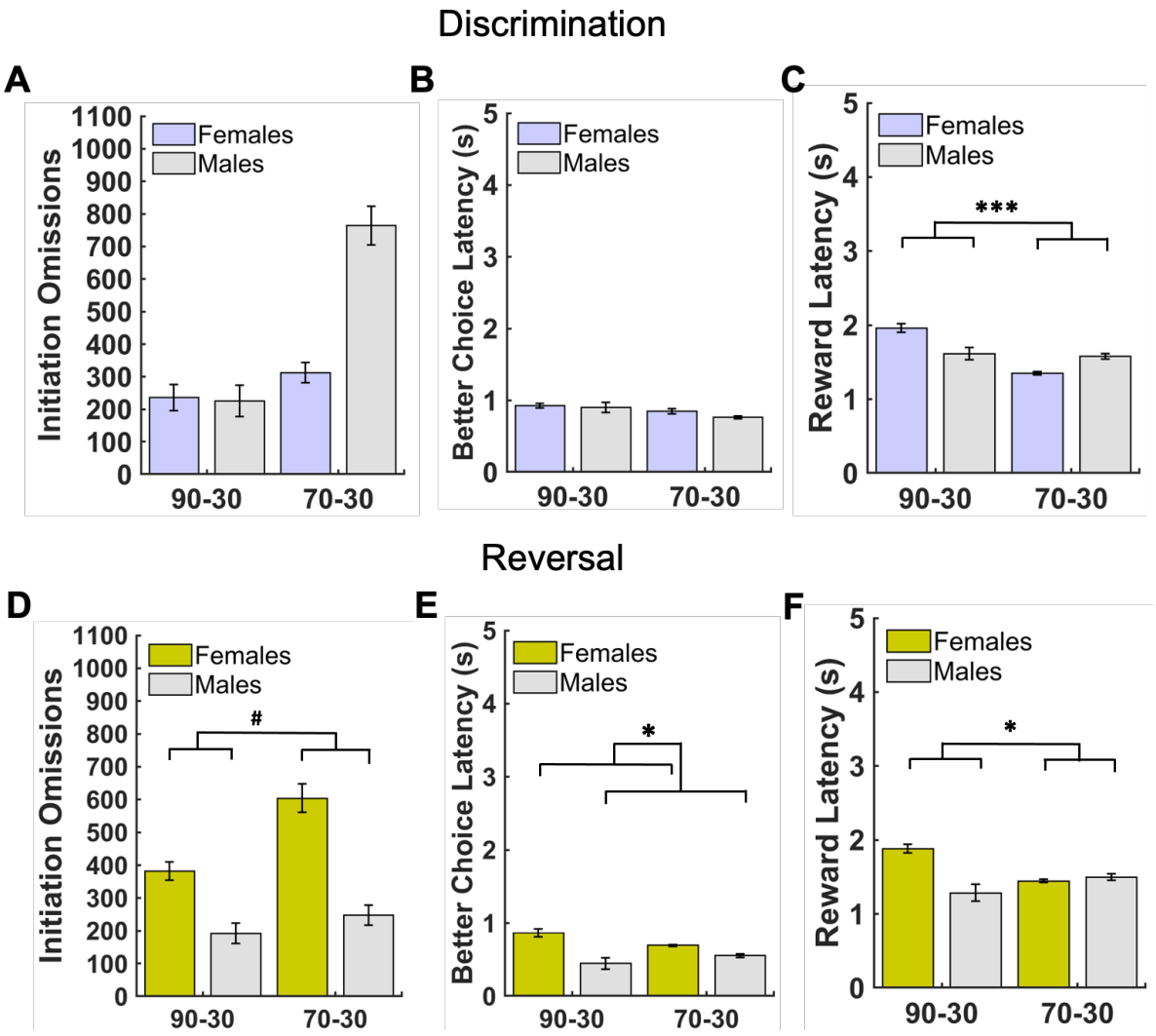

Figure 3-4. Patterns of latencies by sex and reward probability group during discrimination and reversal learning. (A) There were no group or sex differences in initation omissions in discrimination. (B) There were no group or sex differences in better choice latencies in discrimination. (C) There were significant probability group differences in reward collection latencies in discrimination, with the 90-30 reward probability group exhibiting longer latencies than the 70-30 reward probability group. (D) There were no group or sex differences in initation omissions in reversal. (E) There were sex differences in better choice latencies in the reversal phase, with females taking longer to make a choice of the better option than males (with and without controlling for the number discrimination sessions to criterion). (F) There were significant probability group differences in reward collection latencies in the reversal phase, with the 90-30 reward probability group exhibiting longer latencies than the 70-30 reward probability group (with and without controlling for the number of discrimination sessions to criterion). Bars indicate ± SEM $^{\#}p=0.07$, *p ≤0.05, ***p≤0.001

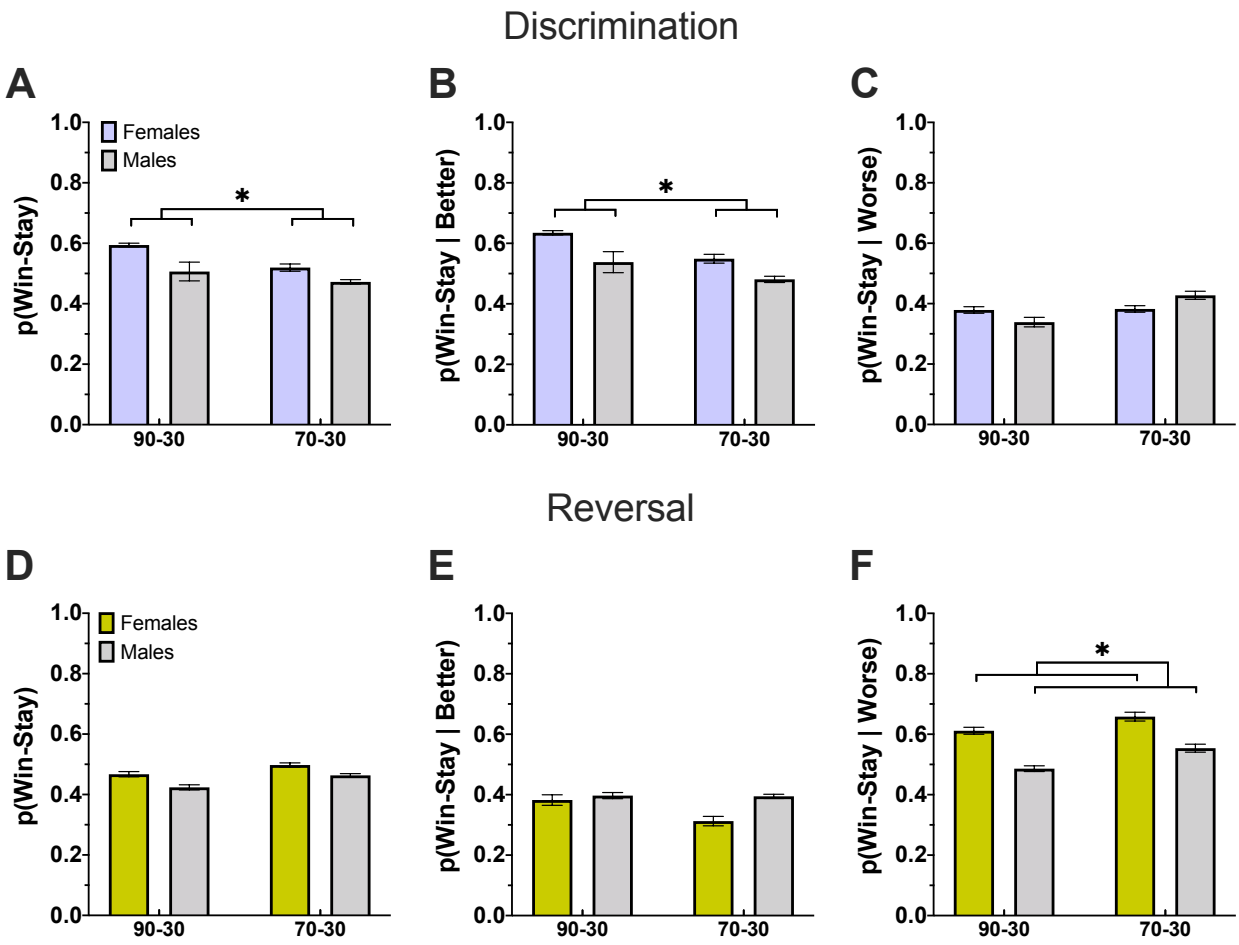

Figure 3-5. Greater overall win-stay and win-stay on the better option in the 90-30 group during discrimination. (**A-C)** Plotted are proportion of win-stay responses overall (A), after choosing the better option (B), and after choosing the worse option (C) during discrimination. Overall win-stay and win-stay on the better option is used more often in the 90-30 group than the 70-30 group. **(D-F)** The same as A-C but for reversal. We find no significant effects on overall win-stay and win-stay on the better option, but do find females are more likely to apply a win-stay strategy after choosing the worse option than males. Bars indicate $\pm$ *SEM* *p $\leq$0.05.

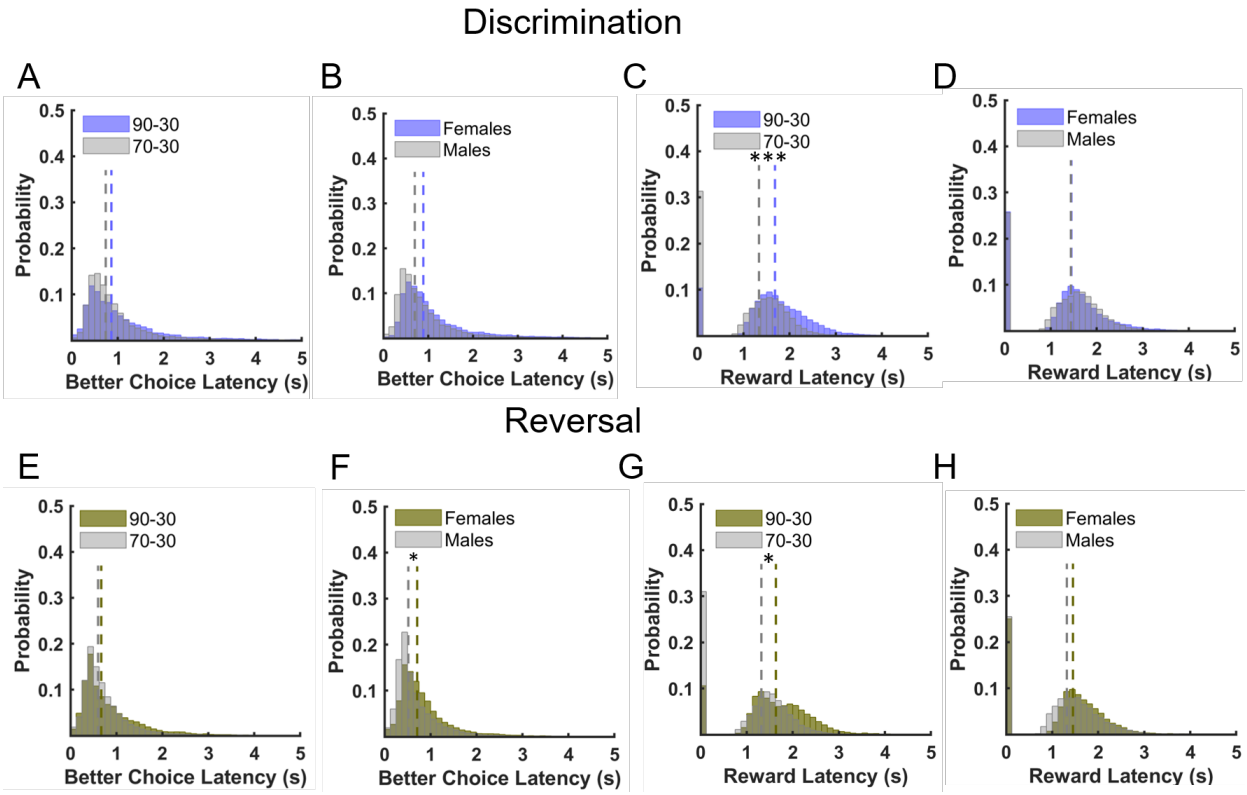

Figure 3-6. 90-30 probability group was consistently slower to collect reward than 70-30 probability group. (A-B) There were no group or sex differences in better choice latencies for discrimination learning. **(C)** The 90-30 group exhibited longer reward collection latencies than the 70-30 group in discrimination learning. **(D)** No significant sex differences in reward collection latencies in discrimination learning. **(E)** No significant group differences in better choice latencies in reversal learning. **(F)** Females exhibited longer latencies for the better option in reversal learning, with or without controlling for discrimination sessions to criterion. **(G)** The 90-30 group exhibited longer reward collection latencies than the 70-30 group in reversal learning, with or without controlling for discrimination sessions to criterion. (**H**) No significant sex differences in reward collection latencies in reversal learning. Dashed lines in histograms of latencies represent group medians. *p ≤0.05, ***p≤0.001

Chapter 4: Sex-dependent contributions of ventrolateral orbitofrontal cortex and basolateral amygdala to learning under uncertainty

ABSTRACT

Reversal learning measures the ability to form flexible associations between choice outcomes with stimuli and actions that precede them. This type of learning is thought to rely on several cortical and subcortical areas, including highly interconnected orbitofrontal cortex (OFC) and basolateral amygdala (BLA), and is often impaired in various neuropsychiatric and substance use disorders. However, unique contributions of these regions to stimulus- and action-based reversal learning have not been systematically compared using a chemogenetic approach and particularly before and after the first reversal that introduces new uncertainty. Here, we examined the roles of ventrolateral OFC (vlOFC) and BLA during reversal learning. Male and female rats were prepared with inhibitory DREADDs targeted in these regions and tested on a series of deterministic and probabilistic reversals during which they learned about stimulus identity or side (left or right) associated with higher reward probability. Using a counterbalanced within-subject design, we inhibited these regions prior to reversal sessions. We measured initial and pre-post reversal changes in accuracy to measure first detection and adjustment to reversals, respectively. We found that inhibition of vlOFC, but not BLA, eliminated detection of stimulus-based reversals. Conversely, both BLA and vlOFC inhibition resulted in significantly slower action-based reversal learning in females, not males, indicating a sex-mediated role for these regions in this type of learning. Learning in females was more impacted in first reversal by vlOFC inhibition than inhibition of BLA, the latter more involved in probabilistic reversal learning. These findings add to mounting evidence of sex-modulated learning flexibility.

INTRODUCTION

Reversal learning, impacted in various neuropsychiatric conditions, measures subjects' ability to form flexible associations between stimuli and actions with outcomes (Dalton et al., 2016; Izquierdo et al., 2013; Schoenbaum et al., 2003) . Reversal learning tasks can also be used to probe learning following expected and unexpected uncertainty in the reward environment (Behrens et al., 2007; Jang et al., 2015; Soltani & Izquierdo, 2019b; Winstanley & Floresco, 2016). For example, after the experience of the first reversal, all others are expected (Jang et al., 2015). Additionally, unexpected uncertainty can be introduced by changes in reward probabilities, after taking the baseline, expected uncertainty into account.

The basolateral amygdala (BLA) is an area of interest in reversal learning due its involvement in value updating (Groman et al., 2019; Janak & Tye, 2015; Tye & Janak, 2007; Wassum & Izquierdo, 2015) and the encoding of both stimulus-based and action-outcome associations typically probed in Pavlovian-to-Instrumental tasks (Corbit & Balleine, 2005; Lichtenberg et al., 2017; Malvaez et al., 2019; Sias et al., 2021). Manipulations of amygdala and specifically BLA have resulted in reversal learning impairments (Churchwell et al., 2009; Groman et al., 2019; Schoenbaum et al., 2003), impaired learning from positive feedback (Costa et al., 2016; Groman et al., 2019), enhanced learning from negative feedback (Izquierdo et al., 2013; Rudebeck & Murray, 2008; Taswell et al., 2021), and even improvements of deficits produced by OFC lesions (Stalnaker et al., 2007). Yet BLA has not been extensively studied in the context of flexible reversal learning of stimuli vs. actions with the exception of a recent lesion study in rhesus macaques (Taswell et al., 2021). BLA has also not been systematically evaluated for its contributions to deterministic vs. probabilistic schedules, with the exception of another lesion study in monkeys (Costa et al., 2016). The idea that BLA encodes changes in the environment in

terms of salience and associability (Roesch et al., 2010) suggests this region may facilitate rapid updating to incorporate new information. The contribution of BLA to reversal learning and its dependence on the nature of the association (i.e., stimulus- vs. action-based), sensory modality (i.e., visual), and type of uncertainty introduced by the task design (i.e., deterministic vs. probabilistic, but also first reversal versus all subsequent reversals) has not been extensively studied using a chemogenetic approach in rats.

In parallel, studies with manipulations in rat OFC in reversal learning have included targeting of the entire ventral surface (Izquierdo, 2017), or systematic comparisons of medial vs. lateral OFC (Hervig et al., 2020; Verharen et al., 2020). Here we examined the role of vlOFC, a subregion not as often probed in reward learning as medial and more (dorso)lateral OFC (cf. Zimmermann et al. (2018)) but also densely-interconnected with BLA (Barreiros, Ishii, et al., 2021; Barreiros, Panayi, et al., 2021). Additionally, unlike almost all previous studies on reversal learning, we included both male and female subjects.

Using a within-subject counterbalanced design, we inactivated these regions prior to reversal sessions and measured both learning and detection/adjustment to reversals. We found that vlOFC, but not BLA, inhibition impaired detection of deterministic and probabilistic stimulus-based, reversals. Conversely, BLA and vlOFC inhibition resulted in significantly slower action-based reversal learning in females, but not males, suggesting a sex-mediated role for these regions. Learning in females was more impacted in first (deterministic) reversal by vlOFC inhibition, and more robustly affected by BLA inhibition in probabilistic reversal. These results suggest similar roles for vlOFC and BLA in flexible action-based learning, but a more specialized role for vlOFC in setting adjustments in stimulus-based learning. Further, fitting choice data with reinforcement learning models indicated the attenuated probabilistic action-based reversal learning deficits were

mediated by a larger memory decay for the unchosen option, especially following vlOFC inhibition. Finally, our findings underscore the importance of including both male and female animals in neuroscience studies, adding to the mounting evidence of sex-modulated learning flexibility (Chen, Ebitz, et al., 2021; Chen, Knep, et al., 2021).

RESULTS

Ex vivo calcium imaging in slices

We performed ex *vivo* Ca^{2+} imaging to confirm the selective action on $CaMKII^{+}$ neuronal excitability in vlOFC and BLA in rats expressing hM4Di DREADD vs. controls expressing mCherry. In BLA, there was no significant effect of CNO (10μM) on Ca^{2+} events for neurons expressing GCaMP6f or GCaMP6f+mCherry (**Figure 4-1C**). A 2-way ANOVA resulted in a significant drug × virus interaction [$F_{(2,324)} = 3.367$, $p = 0.036$], with a selective reduction in the frequency of elicited Ca^{2+} events during CNO only in neurons expressing GCaMP6f+hM4Di (multiple comparison test, p=0.049).

In vlOFC, there was also no significant effect of CNO (10μM) on Ca^{2+} events for neurons expressing GCaMP6f or GCaMP6f+mCherry (**Figure 4-1F**). However, in $CaMKII^{+}$ vlOFC neurons expressing GCaMP6f+hM4Di there was a decrease in the frequency of Ca^{2+} events during CNO application. A 2-way ANOVA revealed a significant drug x virus interaction [$F_{(2,400)} = 8.349$, $p< 0.001$], with multiple comparisons test resulting in decreased Ca^{2+} events in GCaMP6f+hM4Di following CNO ($p = 0.02$), and increased activity in GCaMP6f expressing neurons after CNO ($p = 0.02$).

Discrimination learning: eGFP controls

Mixed-effects GLMs for the discrimination learning phase were conducted for each task separately to establish if there were baseline differences in learning measures between animals infused with eGFP virus in different brain regions. There were no differences between the eGFP groups by target region on learning (i.e., the probability of choosing the correct side) across trials in the action-based task (β_{region} = -0.13, t(2392) = -0.72, p = 0.47), as well as no differences in learning (i.e., probability of choosing the correct visual stimulus) across sessions in the stimulus-based task (β_{region} = 0.10, t(77) = 1.70, p = 0.09). Thus, animals' data were collapsed into a single eGFP virus group for subsequent analyses.

Discrimination learning: hM4Di vs. eGFP

For the action-based task, there were no significant effects of virus or virus interactions for vlOFC vs. eGFP on *probability correct* (β_{virus} = -0.13, t(4792) = -1.02, p = 0.31), with similar findings for the comparison of BLA vs. eGFP (β_{virus} = -0.14, t(4792) =- 1.06, p= 0.29; **Figure 4-2C**). All animals met criterion very quickly (~2 days), thus, we compared trials to reach 75% criterion (i.e., probability of choosing the correct side). Both hM4Di virus groups performed comparably [M±SEM: vlOFC hM4Di (81.1±23.0), BLA hM4Di (84.1±21.7)], whereas the eGFP group met criterion within fewer trials (59.5±18.6), but the difference was not statistically significant [vlOFC hM4Di vs. eGFP: β_{virus} = 54.5, t(25) = 1.34, p = 0.19; BLA hM4Di vs. eGFP: β_{virus} = 54.8, t(25) = 1.38, p = 0.18].

For the stimulus-based task, there were also no significant effects of virus or virus interactions for either vlOFC vs. eGFP on *probability correct* (β_{virus} = -0.06, t(141) = -1.26, p = 0.21), or for BLA vs. eGFP (β_{virus} = -0.05, t(152) = -1.32, p = 0.19; **Figure 4-4C**). The animals on average took approximately ~6 days to meet criterion regardless of virus group [M±SEM: vlOFC

hM4Di (6.1±0.7), BLA hM4Di (6.5±1.2), eGFP (6.9±0.7)]. However, many animals did not meet criterion after a maximum of 10 days of testing (61% of rats).

Given the poorer learning in the stimulus-based task, we evaluated whether this was due to the order of task administered [i.e., Stimulus → Action or Action → Stimulus]. To test whether learning was influenced by task order, we analyzed *probability correct* during initial discrimination learning for the stimulus-based task, which resulted in no effect of task order (β_{order} = 0.03, t(220) = 0.97, p = 0.33), but a significant task order x session interaction ($\beta_{order\ x\ session}$ =-0.03, t(220) = -3.06, p = 0.002). Thus, subsequent analyses were conducted with task order analyzed separately by session, which revealed that animals administered the Action → Stimulus task order exhibited poorer learning across sessions ($\beta_{session}$ = 0.01, t(58) = 2.03, p = 0.05), compared to those administered the Stimulus→ Action task order ($\beta_{session}$ = 0.04, t(162) = 6.14, p < 0.0001). Notably, only 1 BLA hM4Di animal successfully met criterion on stimulus-based learning for the Action → Stimulus order, while no OFC hM4Di animals achieved this.

Accuracy in reversal learning across session and trials

Action-based reversal learning. Mixed-effects GLMs were used to analyze *probability correct*, with trial number, reversal number, drug order, drug, virus, and sex as between-subject factors, trial number, reversal number, and drug as within-subject factors, and individual rat as random factor. GLMs were conducted separately by target region (vlOFC vs. eGFP and BLA vs. eGFP), using the following formula for the full model: *y ~ [1 + trial number * reversal number * virus *drug * drug order * sex + (1 + trial number * reversal number * drug| rat)*].

For the comparison of vlOFC with eGFP, several interactions were found: interaction of sex, virus, drug, and trial number ($\beta_{sex\ x\ virus\ x\ drug\ x\ trial\ number}$ = 0.035, t(19136) = 2.16, p = 0.03), an

interaction of sex, virus, drug, drug order, and trial number ($\beta_{\text{sex x virus x drug order x trial number}}$ = -0.006, t(19136) = -2.32, p = 0.02), as well as an interaction of virus, drug, drug order, reversal number, and trial number ($\beta_{\text{virus x drug x drug order x reversal number x trial number}}$ = -0.001, t(19136) = -2.19, p = 0.03). Due to these interactions, we were justified to look further at individual reversals. For the comparison of vlOFC hM4Di vs eGFP, sex emerged as a significant predictor of R1 probability correct (β_{sex} = -0.279, t(4784) = -2.03, p = 0.04), and thus sex was entered in the model as a covariate: $\gamma \sim [1 + \text{trial number} * \text{reversal number} * \text{virus} * \text{drug order} + \text{sex} + (1 + \text{trial number} * \text{drug} | \text{rat})]$. With sex as a covariate, there was a significant effect of trial number ($\beta_{\text{trial number}}$ = 0.003, t(4791) = 7.08, p = 1.63e^{-12}) and a nonsignificant but trending interaction of virus, drug order, and trial number for R1: ($\beta_{\text{virus x drug order x trial number}}$ = -0.002, t(4791) = -1.80, p = 0.07; **Figure 4-2D**). For probability correct in R3, we also found sex a significant moderator, interacting with drug ($\beta_{\text{sex x drug}}$ = 0.334, t(4784) = 2.18, p = 0.029), and drug x trial number ($\beta_{\text{sex x drug x trial number}}$= -0.003, t(4784) = -2.55, p = 0.01; **Figure 4-2D**). When sex was entered as a covariate in the model, we found a significant effect of trial number as all animals exhibited improvements in probability correct across trials ($\beta_{\text{trial number}}$ = 0.003, t(4791) = 7.73, p = 1.34e^{-14}), with males performing significantly better than females (β_{sex} 0.0841, t(4791) = 2.088, p = 0.04; **Figure 4-3B**), but no other significant interactions with any other factor for action-based reversal learning for vlOFC hM4Di compared to eGFP.

Starting with the full model as above for the comparison of BLA with eGFP, several interactions of virus x drug were observed, including: virus x drug order x trial number, ($\beta_{\text{virus x drug order x trial number}}$ = -0.008, t(18986) = -2.58, p = 0.01), virus x reversal number x drug order x trial number ($\beta_{\text{virus x reversal number x drug order x trial number}}$ = 0.002, t(18986) = 2.59, p = 0.01), and virus x drug x reversal number x drug order x trial number ($\beta_{\text{virus x drug x reversal number x drug order x trial number}}$ = -0.002,

t(18986) = -2.00, p = 0.046). There were also interactions with sex, including: sex x drug x drug order ($\beta_{\text{sex x drug x drug order}}$ = -1.07, t(18986) = -2.72, p = 0.01), and sex x drug x reversal number x drug order ($\beta_{\text{sex x drug x reversal number x drug order}}$ = 0.335, t(18986) = 2.19, p = 0.03). Due to these interactions, we were justified to look further at individual reversals. For BLA hM4Di compared to eGFP in R1 there was also a sex difference (females > males) and an effect of trial number on probability correct (β_{sex} = -0.279, t(4784) = -2.38, p = 0.02; $\beta_{\text{trial number}}$ = 0.002, t(4784) = 3.87, p = $1.0e^{-04}$). With sex as a covariate, there was only a significant effect of trial number for R1 ($\beta_{\text{trial number}}$ = 0.003, t(4791) = 7.17, p = $8.93e^{-13}$). Sex was also a significant moderator of probability correct in R3, with significant interactions of sex x trial number ($\beta_{\text{sex x trial number}}$= 0.002, t(4784) = 2.25, p = 0.03) and sex x trial number x drug ($\beta_{\text{sex x drug x trial number}}$= -0.003, t(4784) = -2.42, p = 0.02). When sex was included as a covariate, there was a significant effect of BLA inhibition on probability correct in probabilistic R3: (GLM: $\beta_{\text{drug*virus}}$ = -0.31, t(4791)= -2.08, p = 0.038). Bonferroni-corrected post-hoc comparisons revealed an effect of CNO in hM4Di (p < 0.01), not in eGFP (p = 1.0), with both males and females exhibit attenuated learning of probabilistic R3 following BLA inhibition.

Given the sex x drug interactions observed in both hM4Di groups, *probability correct* was also analyzed separately for hM4Di males, eGFP males, hM4Di females, and eGFP females using the following formula: $\gamma \sim [1 + drug + (1 + drug| rat)]$. We found a significant effect of drug for R1 (β_{drug}= -0.24, t(1348) = -2.49, Bonferroni corrected p = 0.03) and R3 (β_{drug}= -0.11, t(1348) = -3.09, Bonferroni corrected p = 0.004) for vlOFC hM4Di females and significant effect of drug for R3 for BLA hM4Di females (β_{drug}= -0.27, t(1348) = -2.97, Bonferroni corrected p = 0.006). There was no significant effect of drug in eGFP females, eGFP males, or hM4Di males across regions and reversals (**Figure 4-3**).

Stimulus-based reversal learning. In contrast to the acquisition curves that demonstrated learning of the initial visual discrimination (**Figure 4-4C**), all animals exhibited difficulty with stimulus-based reversal learning, rarely achieving above 60% after 10 sessions (**Figure 4-8**), similar to recent reports (Harris, Aguirre et al., 2021; Ye et al., 2023). Here, due to several non-learners, we adhered to the criterion of rats reaching greater than a 50% running window average for the last 100 trials in discrimination, for inclusion in subsequent reversal learning analyses. The following numbers did not meet this criterion and were excluded from these groups: 0 of 13 vlOFC hM4Di, 6 of 15 BLA hM4Di, and 6 of 17 eGFP. As above, GLMs were conducted separately for accuracy (*probability correct*) by target region comparison, but with session instead of trial number as a within-subject factor. Thus, the GLM formula was as follows: $\gamma \sim [1 + session * reversal\ number * virus * drug * drug\ order * sex + (1 + session * reversal\ number * drug | rat)]$.

For vlOFC hM4Di comparison with eGFP, we observed several interactions including: session, drug, and virus ($\beta_{session\ x\ drug\ x\ virus} = -0.096$, $t(536) = -2.83$, $p = 4.81e^{-03}$), session, drug, drug order, and virus ($\beta_{session\ x\ drug\ x\ drug\ order\ x\ virus} = 0.146$, $t(536) = 2.62$, $p = 9.17e^{-03}$) and drug, drug order, and reversal number ($\beta_{drug\ x\ drug\ order\ x\ reversal\ number} = 0.107$, $t(536) = 2.18$, $p = 0.03$). We also observed a sex x reversal number interaction ($\beta_{sex\ x\ reversal\ number} = -0.084$, $t(536) = -2.50$, $p = 0.01$). Aside from a significant effect of session in R1 ($\beta_{session} = 0.03$, $t(134) = 2.67$, $p = 8.55e^{-03}$) there were no significant predictors of learning with follow-up GLM analyses with sex in the model. With sex as a covariate in the model there was a significant effect of session for R1 and R3 (R1: $\beta_{session} = 0.026$, $t(141) = 3.80$, $p = 2.16e^{-04}$; R3: $\beta_{session} = 0.018$, $t(141) = 2.92$, $p = 4.11e^{-03}$), and a session x virus interaction in R3 ($\beta_{session\ x\ virus} = -0.020$, $t(141) = -2.07$, $p = 0.04$). Bonferroni-corrected post-hoc comparisons revealed an effect of session in eGFP ($p < 0.01$), but not in hM4Di ($p = 0.45$), indicating that only the eGFP group improved across session in R3 (**Figure 4-8**).

For BLA hM4Di compared to eGFP, several interactions were observed including interactions of session, drug, and virus ($\beta_{\text{session x drug x virus}} = -0.087$, $t(575) = -2.25$, $p = 0.02$), drug, drug order, and reversal number ($\beta_{\text{drug x drug order x reversal number}} = 0.107$, $t(575) = 2.29$, $p = 0.02$), and sex and reversal number ($\beta_{\text{sex x reversal number}} = -0.084$, $t(575) = -2.15$, $p = 0.03$). However, none of the post-hoc GLM analyses resulted in significant predictors of learning with the exception of session for R1 and R3 (R1: $\beta_{\text{session}} = 0.026$, $t(151) = 2.35$, $p = 0.02$; R3: $\beta_{\text{session}} = 0.018$, $t(151) = 2.67$, $p = 8.43e^{-03}$) when sex was included as a covariate in the model.

Due to the slow stimulus-based reversal learning, we next assessed *probability correct* around reversals (three sessions before and after reversals for stimulus-based learning, and one session before and after reversals for action-based learning) to test for detection and adjustments to reversals.

Estimated model parameters from the reinforcement learning models.

Refer to Appendix C.

Accuracy around reversals: Reversal detection

Stimulus-based reversal learning. We analyzed accuracy (*probability correct*) around reversals, as overall stimulus-based reversal learning was modest. ANOVAs with virus and drug order as between subject factors were conducted on the mean change in accuracy between one reversal and the next. vlOFC hM4Di was significantly different from eGFP for R1-to-R2 [$F_{(1,24)} = 9.49$, $p < 0.01$] and R2-to-R3 [$F_{(1,24)} = 10.1$, $p < 0.01$], but not R3-to-R4 [$F_{(1,24)} = 2.61$, $p = 0.12$], **Figure 4-4D**. In contrast, BLA hM4Di was not significantly different from eGFP on changes in accuracy around any of the reversals.

Action-based reversal learning. We also assessed accuracy (*probability correct*) around reversals for action-based reversal learning. As above, ANOVAs with virus and drug order as

between subject factors were conducted on the mean accuracy change between one reversal and the next. Other than confirming the probabilistic reversal learning (R3) impairment for BLA hM4Di (**Figure 4-2**), there were no significant effects of virus groups on accuracy changes on any other reversal transition in action-based reversal learning (**Figure 4-6).**

In summary, our results indicate that both BLA and vlOFC are required for learning action-based reversals in females. Conversely, vlOFC, but not BLA, is necessary for detecting stimulus-based, not action-based, reversals.

Win-Stay, Lose-Switch Strategies around reversals

Stimulus-based reversal learning. We also analyzed adaptive response strategies (*Win-Stay* and *Lose-Switch*) around reversals. ANOVAs with virus and drug order as between subject factors were conducted on mean Win-Stay or Lose-Switch between one reversal and the next. vlOFC hM4Di was significantly different from eGFP for *Win-Stay* R1-to-R2 [$F_{(1,24)}$ =9.91, $p < 0.01$] and R2-to-R3 [$F_{(1,24)}$= 11.61, $p < 0.01$], but not R3-to-R4 [$F_{(1,24)}$= 1.49, $p = 0.24$]. Similarly, vlOFC hM4Di differed from eGFP for *Lose-Switch* in R1-to-R2 [$F_{(1,24)}$=6.00, p=0.023] and R2-to-R3 [$F_{(1,24)}$=5.00, p=0.036], but not R3-to-R4 [$F_{(1,24)}$=0.08, p=0.78], **Figure 4-5**.

In contrast, BLA hM4Di was not significantly different from eGFP on changes in *Win-Stay* or *Lose-Switch* strategies around any of the reversals. Therefore, the results for these adaptive strategies reflect an identical pattern to that observed for *probability correct* for both vlOFC and BLA hM4Di, above.

Performance measures in reversal learning

We analyzed other performance measures including *initiation, choice,* and *reward latencies*, as proxies for attention, deliberation, and motivation, respectively (Harris, Aguirre et al., 2021; Ye et al., 2023). Given that the significant effects in accuracy were observed only in R1 and R3, we elected to analyze latency measures during those reversals only, with sex added as a covariate in the GLM models. In this case, a single measure per animal per reversal number was obtained: the GLM formula was as follows for R1 and R3: $\gamma \sim [1 + virus * drug + sex + (1 + drug| rat)]$.

Action-based reversal learning. vlOFC comparisons with eGFP controls did not yield any significant effects for R1; however, a significant drug effect emerged in R3, such that VEH-treated animals exhibited longer *initiation latencies* than when treated with CNO, irrespective of virus group (β_{virus} = -0.524, t(27) = -2.06, p = 0.049). Females tended to exhibit longer latencies than males, although this was only trending toward significance (β_{sex} = -0.353, t(27) = -1.96, p = 0.06).

When comparing BLA hM4Di vs. eGFP, we found a significant effect of virus in the full model (β_{virus} = 1.44, t(93) = 2.07, p = 0.039), with BLA hM4Di animals exhibiting longer initiation latencies than eGFP controls. We also observed a significant effect of virus for both *incorrect choice* (R1: β_{virus} = 0.38, t(27) = 4.68 p < 0.0001; R3: β_{virus} = 0.28, t(28) = 3.40, p = 0.002) and *reward latencies* (R1: β_{virus} = 0.34, t(30) = 3.78 p = 0.0007; R3: β_{virus} = 0.22, t(30) = 2.31, p = 0.03), such that BLA hM4Di exhibited longer latencies compared to eGFP (**Figure 4-7B**). This effect was not observed when comparing vlOFC hM4DI vs. eGFP (**Figure 4-7A**).

Stimulus-based reversal learning. We similarly analyzed performance measures in stimulus-based reversal learning during R1 and R3, with sex added as a covariate to the model. We found a robust sex difference, with females committing more *initiation omissions* than males, irrespective of virus or drug, across reversals (vlOFC vs. eGFP: β_{sex} = -88.52, t(80) = -2.40 p =

0.02; BLA vs. eGFP: β_{sex} = -80.38, t(89) = -2.14, p = 0.04; **Figures 4-9A, 4-9B**). For *initiation latencies,* vlOFC hM4Di vs. eGFP analysis yielded an overall significant effect of sex as well (β_{sex} = -3.23, t(81) = -3.72, p = 0.0004), such that females exhibited longer initiation latencies than males across reversals (**Figures 4-9C, 4-9D**). BLA hM4Di vs. eGFP comparisons also yielded a significant effect of sex (β_{sex} = -3.26, t(88) = -3.88, p = 0.0002), however, there was also a significant interaction between reversal number, virus, drug, drug order, and sex ($\beta_{\text{reversal number x virus x drug x drug order x sex}}$ = 1.46, t(88) = 2.13, p = 0.036), which justified analysis of each reversal number separately. Subsequent analyses revealed a significant effect of sex for both R1(β_{sex} = -1.58, t(24) = -3.40, p = 0.002), and R3 (β_{sex} = -0.67, t(24) = -2.42, p = 0.024), with females exhibiting longer initiation latencies than males, irrespective of virus or drug (**Figure 4-9D**).

In summary, across all reversals BLA hM4Di animals exhibited slower deliberation speed during incorrect choices, and took longer to collect reward compared to eGFP controls in the action-based task. In contrast, sex emerged as a strong predictor of performance measures, but not learning, in the stimulus-based task, as it accounted for much of the variance in initiation omissions and latencies.

DISCUSSION

We used a chemogenetic approach to transiently inactivate neurons in either vlOFC or BLA to assess how these regions are involved in different aspects of reversal learning. Although the role of OFC in reversal learning has been instantiated in different paradigms using visual stimuli and cues (Izquierdo et al., 2013; Piantadosi et al., 2018; Hervig et al., 2020; Alsio et al., 2021) as well as olfactory ones (Schoenbaum et al., 2003; Kim and Ragozzino, 2005), several groups also report a strong role for OFC in action (spatial)-based reversal learning (Dalton et al., 2016; Groman et

al., 2019; Verharen et al., 2020). Almost all of these reversal learning investigations have involved irreversible lesions or baclofen/muscimol inactivations of OFC. Testing both types with a chemogenetic approach, here we found that different aspects of both action- and stimulus-based reversal learning rely on vlOFC.

In parallel, the specific role of BLA in stimulus- vs. action-based reversal learning is poorly understood given mixed results (Schoenbaum et al., 2003; Izquierdo and Murray, 2004; Churchwell et al., 2009; Hervig et al., 2020). Recent studies suggest amygdala may be involved in both types of learning (Taswell et al., 2021; Keefer and Petrovich, 2022) as BLA activity is modulated by violations in reward expectations generally, which are not association-specific (Roesch et al., 2012). To probe this, we tested animals on both stimulus- and action-based tasks and found that BLA is more selectively involved in action-based reversal learning.

As additional motivations for the present study, several reports suggest that neural recruitment in reversal learning may depend on certainty of rewards (Boulougouris et al., 2007; Boulougouris and Robbins, 2009; Ward et al., 2015; Costa et al., 2016; Dalton et al., 2016; Piantadosi et al., 2018; Verharen et al., 2020). To further understand this, we tested animals on both deterministic (100/0) and probabilistic reversals (90/10). We found vlOFC to be involved in the detection of stimulus-based reversals and initial learning of both deterministic and probabilistic learning across stimuli and actions, whereas BLA was more selectively involved in probabilistic reversal learning of actions.

Finally, due to the sparsity of research probing sex differences in flexible learning and decision making (Orsini and Setlow, 2017; Grissom and Reyes, 2019; Orsini et al., 2022; Cox et al., 2023), where an overwhelming number of reversal learning studies include only males (Schoenbaum et al., 2003; Izquierdo et al., 2013; Dalton et al., 2016; Groman et al., 2019; Hervig

et al., 2020; Verharen et al., 2020), we included both male and female rats here. We found sex-dependent contributions of both vlOFC and BLA in action-based reversal learning. We elaborate on these findings within the context of the existing literature below.

Similar recruitment of BLA and vlOFC during action-based reversal learning

All animals learned to flexibly adjust their responses following deterministic and probabilistic reversals, indicating successful remapping of reward contingencies as accuracy increased across trials. Importantly, we found no effect of CNO in eGFP animals, suggesting that it was activation of hM4Di receptors in BLA and vlOFC that were crucial to any impairments observed. After considering sex as a covariate in our analyses, we determined that vlOFC was necessary for learning of first deterministic (R1) and probabilistic reversal (R3). This is consistent with findings following pharmacological inactivations or lesions of OFC (Boulougouris et al., 2007; Boulougouris and Robbins, 2009; Dalton et al., 2016; Piantadosi et al., 2018; Verharen et al., 2020).

BLA inhibition was not expected to impair deterministic reversal learning as it is thought to be mostly recruited when there is some level of uncertainty, e.g., probabilistic outcomes (Roesch et al., 2012). Amygdala-lesioned monkeys are also impaired on action(spatial)-based probabilistic reversal learning, exhibiting decreased probability of choosing the better option, and increased switching behavior following negative outcomes (Taswell et al., 2021). BLA may indeed be critical in generating prediction error signals following changes in reward associations (Esber et al., 2012; Roesch et al., 2012; Iordanova et al., 2021), with particular involvement in detecting unexpected upshifts or downshifts in value (Roesch et al., 2010; Stolyarova and Izquierdo, 2017). Our finding of attenuated learning of probabilistic reversal R3 suggests it is the misleading

feedback that most engages BLA. Future investigations should probe the role of BLA in initial learning of probabilistic outcomes, without reversal experience.

vlOFC, but not BLA, is necessary for detection of reversals in stimulus-based learning

As described above, unlike the ease of action-based reversal learning, rats exhibited difficulty learning reversals of stimulus-reward contingencies, as previously reported (Harris, Aguirre et al., 2021). Thus, instead of examining acquisition curves which reached asymptote slightly above chance, we elected to study detection and adjustment to reversals by comparing accuracy and strategy prior to and after a reversal occurred. Furthermore, this enabled assessment about whether prior inhibition affected future detection of reversals and whether this varied by transition type [i.e., between deterministic reversals (R1→R2), deterministic and probabilistic reversal (R2→R3), or probabilistic reversals (R3→R4)]. We found that vlOFC, but not BLA, inhibition produced a failure in detecting first deterministic and first probabilistic reversal. This pattern was not observed in animals that received VEH during the first deterministic reversal, suggesting vlOFC needs to be "online" when experiencing a reversal for the very first time as this determines how flexibly animals respond to future reversals. Employment of adaptive strategies matched this effect, such that vlOFC-inhibited animals did not employ either win-stay or lose-switch strategies after the first reversal. That vlOFC inhibition did not impair the ability to detect transitions between probabilistic reversals (R3→R4) supports the idea that other brain regions may be recruited when the probabilistic reward contingencies have already been established. The role of OFC in establishing an "expected uncertainty" (Soltani and Izquierdo, 2019) has been instantiated experimentally in several recent studies using different methodologies (Namboodiri et al., 2019; Namboodiri et al., 2021; Jenni et al., 2022), and we add detection and adjustment to

stimulus-based reversals to this evidence. In contrast, BLA inhibition did not result in any impairment in the ability to detect and flexibly adjust to reversals, regardless of whether they were deterministic or probabilistic.

Females exhibited poorer action-based reversal learning following vlOFC and BLA inhibition

We found a significant sex-dependent effect of vlOFC inhibition for both the first deterministic (R1) and the first probabilistic reversal (R3), and a similar effect in females following BLA inhibition for R3. Importantly, although BLA hM4Di males also exhibited the general pattern of attenuated learning of R3, the effect of BLA inhibition was largely driven by females. These findings were unexpected as we did not anticipate sex differences in the recruitment of brain regions involved in action-based reversal learning, mostly due to the lack of studies to date investigating the effect of sex. Nonetheless, there is evidence that OFC is differentially activated in males and females during risky decisions such that activity in the lateral OFC is inversely correlated with the proportion of advantageous choices in female, but not in male, rats (van Hasselt et al., 2012), with similar effects found in humans (Bolla et al., 2004) and in amygdala (Dreher et al., 2007).

Unfortunately, the potential of estrous-driven behavioral variation in females has commonly been used as a rationale for excluding female rodents in behavioral neuroscience research (Beery and Zucker, 2011; McCarthy et al., 2012; Shansky and Murphy, 2021). However, our findings, and other recent studies of cortical circuits exhibiting sex-mediated influences on reward-motivated behavior (Cox et al., 2023) should instead encourage the inclusion of both sexes in experiments as clear patterns emerge. Additionally, the learning impairment observed in females following vlOFC and BLA inhibition may not be primarily due to fluctuations in hormone levels,

but rather may reflect differential adoption of strategies between sexes (Grissom and Reyes, 2019; Chen et al., 2021a; Chen et al., 2021b). Consistent with this view, our results based on estimated RL model parameters suggests differential mechanisms for adjustment to reversals between males and females. Specifically, we found some evidence for differential effects on γd, where the decay rate for unchosen action values was greater for females than males. This is consistent with a previous study which also observed similar disruption in retention of action values after ablating OFC neurons projecting to BLA (Groman et al., 2019). Given that study involved only male rats and involved a stronger manipulation than our chemogenetic approach (i.e., one that caused pathway-specific neuronal apoptosis), it is plausible their observed effect on γd would have been different between females and males.

Stimulus-based vs. action-based learning

Interestingly, we discovered task order to be significant in rats' ability to learn to discriminate stimuli: stimulus→ action was learned much more readily than action → stimulus. This can be explained by noting that rats are heavily biased to acquire spatial associations (Wright et al., 2019), and reinforcing this already-strong learning likely inhibits the ability to learn associations where spatial information should be ignored. In contrast, nonhuman primates are able to quickly transition between "what" vs. "where" blocks of trials (Rothenhoefer et al., 2017; Taswell et al., 2021). Nonetheless, learning both types of associations is crucial for flexibility required in naturalistic environments and thus, it is important to examine how stimulus-based and action-based learning systems interact with each other (Soltani & Koechlin, 2022). Moreover, although the role of OFC in stimulus- or cue-based reversal learning has been probed using

olfactory and visual stimuli, more viral-mediated approaches employing targeted chemogenetic and optogenetic manipulations across sensory modalities in both males and females are warranted.

Conclusion

The present results suggest similar roles for vlOFC and BLA in flexible action-based learning, but a more specialized role for vlOFC in setting adjustments in stimulus-based learning. Additionally, our findings underscore the importance of including both male and female animals in behavioral neuroscience studies, adding to the mounting evidence of sex-modulated learning flexibility (Chen, Ebitz, et al., 2021; Chen, Knep, et al., 2021).

METHODS

Subjects

Animals for behavioral experiments were adult (N=56, n=25 females; 52 for behavior, 4 for ex-vivo imaging) Long-Evans rats (Charles River Laboratories) average age post-natal-day (PND) 65 at the time of order, with a 280g body weight minimum for males and 240g body weight minimum for females at the time of surgery and the start of the experiment. Rats were approximately PND 100 (emerging adulthood; Ghasemi et al. (2021)) when behavioral testing commenced. Before any treatment, all rats underwent a 3-day acclimation period during which they were pair-housed and given food and water *ad libitum*. During that time, they remained in their homecage with no experimenter interference. Following this 3-day acclimation period, animals were handled for 10 min per animal for 5 consecutive days. During the handling period, the animals were also provided food and water *ad libitum*. After the handling period, animals were individually-housed under standard housing conditions (room temperature 22–24° C) with a

standard 12 h light/dark cycle (lights on at 6am). Animals were then surgerized and tested on discrimination and reversal learning 1-week post-surgery. At the point of reversal, they were beyond the 3-week expression time for DREADDs.

A separate group of Long-Evans rats (N=4, all males) were used for validation of effectiveness of DREADDs in BLA and vlOFC in slice, using ex-vivo calcium imaging procedures. All procedures were conducted in accordance to the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and with the approval of the Chancellor's Animal Research Committee at the University of California, Los Angeles.

Surgery

Viral Constructs

Rats were singly-housed and remained in homecages for 4 weeks prior to testing while the inhibitory hM4Di DREADDs expressed in BLA (n=16, 9 females), vlOFC (n=19, 9 females), or eGFP control virus (n=17, 7 females) in these regions. An adeno-associated virus AAV8 driving the hM4Di-mCherry sequence under the CaMKIIa promoter was used to express DREADDs bilaterally in BLA neurons (0.1 µl, AP = -2.5; ML= ±5; DV = −7.8 and 0.2 µl, AP= -2.5; ML= ±5; DV= −8.1, from bregma at a rate of 0.1 µl/min; AAV8-CaMKIIa-hM4D(Gi)-mCherry, packaged by Addgene, viral prep #50477-AAV8). In other animals, this same virus (AAV8-CaMKIIa-hM4Di-mCherry, Addgene) was bilaterally infused into two sites in vlOFC (0.2 µl, AP = +3.7; ML= ±2.5; DV = −4.6 and 0.15 µl, AP= 4; ML= ±2.5; DV= −4.4, from bregma at a rate of 0.1 µl/min). A virus lacking the hM4Di DREADD gene and only containing the green fluorescent tag eGFP (AAV8-CaMKIIa-EGFP, packaged by Addgene) was also infused bilaterally into either

BLA (n=7), vlOFC (n=5), or anterior cingulate cortex [(n=5); 0.3 µl, AP = +3.7; ML= ±2.5; DV = −4.6, rate of 0.1 µl/min] as null virus controls. Our vlOFC targeting is most similar to infusion sites reported previously by Dalton et al. (2016), and 0.7 mm more medial than others (Costa et al., 2023). In rats used for ex vivo calcium imaging, the same target 166 regions were infused with either GCaMP6f (AAV9-CaMKIIa-GCaMP6f, Addgene), a 1:1 combination of GCaMP6f+mCherry (AAV8-CamKIIa-mCherry, Vector BioLabs, #VB1947), or a 1:1 combination of GCaMP6f+hM4Di-mCherry (same as used for behavior, AAV8-CaMKIIa-hM4Di-mCherry, Addgene).

Surgical Procedure

Infusions of DREADD or eGFP control virus were performed using aseptic stereotaxic techniques under isoflurane gas (1-5% in O_2) anesthesia prior to any behavioral testing experience. Before surgeries were completed, all animals were administered 5mg/kg s.c. carprofen (NADA #141–199, Pfizer, Inc., Drug Labeler Code: 000069) and 1cc saline. After being placed in the stereotaxic apparatus (David Kopf; model 306041), the scalp was incised and retracted. The skull was leveled +2mm and -2mm A-P to ensure that bregma and lambda were in the same horizontal plane. Small burr holes were drilled in the skull above the infusion target. Virus was bilaterally infused at a rate of 0.01 µl per minute in target regions (coordinates above). After each infusion, 5 min elapsed before exiting the brain.

Histology

At the end of the experiment, rats were euthanized with an overdose of Euthasol (Euthasol, 0.8 mL, 390 mg/mL pentobarbital, 50 mg/mL phenytoin; Virbac, Fort Worth, TX), were transcardially perfused, and their brains removed for histological processing. Brains were fixed in

10% buffered formalin acetate for 24 h followed by 30% sucrose for 5 days. To visualize hM4Di-mCherry and eGFP expression in BLA or vlOFC cell bodies, free-floating 40-µm coronal sections were mounted onto slides and cover-slipped with mounting medium for DAPI. Slices were visualized using a BZ-X710 microscope (Keyence, Itasca, IL), and analyzed with BZ-X Viewer and analysis software.

Reconstructions of viral expressions of hM4Di (magenta) and green fluorescent protein, eGFP (green) across the AP plane (**Figures 4-1 B-E**) were conducted using Photoshop and Illustrator (Adobe, Inc.) by individuals blind to condition. Two raters then independently used imageJ (U. S. National Institutes of Health, Bethesda, Maryland, USA) to trace and quantify pixels at AP +3.7 (vlOFC) and AP -2.8 (BLA) for each animal. Three measures were obtained per hemisphere and were highly correlated (Spearman rank correlation: $r=0.93$, $p<0.01$). Only subjects with bilateral expression were included in the analysis of behavior. There were no differences in expression level between males and females for pixel count reconstructions [$F(1,12)= 0.32$, $p=0.58$], but there was a significant difference between target brain region (vlOFC, BLA) in the hM4Di virus group [$F(1,12)= 9.71$, $p=0.009$]; the latter was expected given the larger infusion volumes in vlOFC.

Food Restriction

Five days prior to any behavioral testing, rats were placed on food restriction with females on average maintained on 10-12 grams/ day and males given 12-14 grams/ day of chow. Food restriction level remained unchanged throughout behavioral testing, provided animals completed testing sessions. Water remained freely available in the home cage. Animals were weighed every other day and monitored closely to not fall below 85% of their maximum, free-feeding weight.

Drug administration

Inhibition of vlOFC or BLA was achieved by systemic administration of clozapine-N-oxide, CNO (i.p., 3mg/kg in 95% saline, 5% DMSO) in animals with DREADDs. Rats with eGFP in these regions underwent identical drug treatment. Rats were randomly assigned to drug treatment groups, irrespective of performance in pretraining. CNO was administered only during reversal learning, 30 min prior to behavioral testing. We followed previous work on timing and dose of systemic CNO (Hart et al., 2020; Stolyarova et al., 2019) and considering the long duration of test sessions. To control for nonspecific effects of injections and handling stress, we also injected animals with VEH. To increase power and decrease the number of animals used in experiments, we used a within-subject design for assessing the effects of CNO, with all rats receiving CNO and VEH injections in counterbalanced order. Thus, for drug administration if a rat received CNO on the 1st reversal, it was administered VEH on the 2nd reversal, or vice versa.

Behavioral Testing

Pretraining. Behavioral testing was conducted in operant conditioning chambers outfitted with an LCD touchscreen opposing the sugar pellet dispenser. All chamber equipment was controlled by customized ABET II TOUCH software.

The pretraining protocol, adapted from established procedures (Stolyarova & Izquierdo, 2017), consisted of a series of phases: Habituation, Initiation Touch to Center Training (ITCT), Immediate Reward Training (IMT), designed to train rats to nosepoke, initiate a trial, and select a stimulus to obtain a reward (i.e., sucrose pellet). Pretraining stages have been reported in detail elsewhere (Stolyarova et al., 2019). For habituation pretraining, the criterion for advancement was collection of all 5 sucrose pellets. For ITCT, the criterion to the next stage was set to 60 rewards

consumed in 45 min. The criterion for IMT was set to 60 rewards consumed in 45 min across two consecutive days. After completion of all pretraining schedules, rats were advanced to the discrimination (initial) phase of either the action- or stimulus-based reversal learning task, with the task order counterbalanced (**Figure 4-2 A-B** or **Figure 4-3 A-B**). A subset of animals was tested first on the action-based task (10 vlOFC hM4Di, 9 BLA hM4Di), while others were tested first on the stimulus-based task (9 vlOFC hM4Di, 7 BLA hM4Di, 17 eGFP). Three vlOFC rats completed only the stimulus-based task.

Action-based deterministic discrimination learning. After completion of either all pretraining schedules or all four reversals of the stimulus-based task, rats were advanced to the discrimination (initial) phase of the action-based task (**Figure 4-2A**). Rats were required to initiate a trial by touching the white graphic stimulus in the center screen (displayed for 40 s), and after initiation rats would be presented with two identical stimuli (i.e., fan or marble) on the left and right side of the screen (displayed for 60 s), that they were required to nosepoke as either the correct spatial side ($p_R(B)= 1$; rewarded with one sucrose pellet) or incorrect spatial side ($p_R(W)=0$). Thus, rats were required to ignore the properties of the stimuli and determine the better rewarded side. The criterion was set to 60 or more rewards consumed and selection of the correct option in 75% of the trials or higher during a 60 min session across two consecutive days. After reaching the criterion for the discrimination phase, the rats were advanced to the reversal phase beginning on the next session. Animals were not administered either CNO or VEH injections during discrimination learning.

Action-based reversal learning. After the discrimination phase, the rats advanced to the reversal phase. Before a reversal learning session, rats were injected intraperitoneally with either 3 mg/kg of clozapine-N-oxide (CNO), or a saline vehicle (VEH) control 30 min prior to each

reversal testing session. The side previously associated with the p_R (B)=1 probability (i.e., correct option), was now associated with a p_R(W)=0 probability of being rewarded (i.e., incorrect option), and vice versa. The criterion was the same as the deterministic discrimination phase. After reaching the criterion for the first deterministic reversal phase (i.e., R1), the rats advanced to the second deterministic reversal phase (i.e., R2) beginning on the next session. Rats that had previously received VEH during the first reversal would now receive CNO injections, and vice versa.

After completing both action-based deterministic reversal learning, rats advanced to the first probabilistic reversal learning phase (i.e., reversal 3, R3). Rats underwent the same injection procedure as the prior reversals. However, the spatial side (i.e., left or right) previously associated with p_R(C)=1 probability (i.e., correct option), was now associated with a p_R(W)=0.1 probability of being rewarded (i.e., worse option), whereas the spatial side previously associated with the p_R(I)=0.0 probability (i.e., incorrect option), was now associated with a p_R(B)=0.9 probability (i.e., better option). The criterion was the same as the previous deterministic reversal learning phases. After reaching the criterion for the first probabilistic reversal learning phase (i.e., reversal 3, R3), rats were advanced to the second probabilistic reversal phase (i.e., reversal 4, R4) beginning on the next testing day, where the probabilities would be reversed once again. Rats that had previously received VEH during the first probabilistic reversal now received CNO injections, and vice versa.

Stimulus-based deterministic discrimination learning. After completion of all pretraining schedules (or all reversals of the action-based task), rats were advanced to the discrimination (initial) phase of learning in which they would initiate a trial by touching a white graphic stimulus in the center screen (displayed for 40 s), and choose between two different visual stimuli pseudorandomly presented on the left and right side of the screen (**Figure 4-3A**). Stimuli were displayed for 60 s each, randomly assigned as the correct or incorrect stimulus: p_R(B)= 1.0 (i.e.,

correct option) or $p_R(W)=0.0$ (i.e., incorrect option). If a trial was not initiated within 40 s, it was scored as an initiation omission. If a stimulus was not chosen, it was scored as a choice omission, and a 10 s ITI ensued. If a trial was not rewarded, a 5 s time-out would follow, subsequently followed by a 10 s ITI. Finally, if a trial was rewarded, a 10 s ITI would follow after the reward was collected. A criterion was set to 60 or more rewards consumed and selection of the correct option in 75% of the trials or higher during a 60 min session across two consecutive days. After reaching the criterion for the discrimination phase or if they were unable to achieve criterion after 10 days, rats were advanced to the reversal phase beginning on the next session. Animals were not given either CNO or VEH injections during discrimination learning.

Stimulus-based reversal learning. After the discrimination phase, rats advanced to the first deterministic reversal learning phase (i.e., reversal 1, R1) where they were required to remap stimulus-reward contingencies. As above, before a reversal learning session, rats were injected intraperitoneally with either 3 mg/kg of CNO, or a saline VEH control, 30 min prior to each reversal testing session. The criterion was the same as discrimination learning. After reaching the criterion for the first reversal phase or if they were unable to achieve criterion after 10 days, the rats were advanced to the second deterministic reversal phase (i.e., reversal 2, R2) beginning on the next testing day, where the reward contingencies were reversed once again. Rats that had previously received VEH during the first reversal now received CNO injections, and vice versa.

After completing both deterministic reversal learning phases, rats advanced to the first probabilistic reversal learning phase (i.e., reversal 3, R3). The injection procedure remained the same as prior reversals. However, the visual stimulus previously associated with the $p_R(C)=1$ probability (i.e., correct option), would now be associated with a $p_R(W)=0.1$ probability of being rewarded (i.e., worse option), whereas the stimulus previously associated with the $p_R(I)=0$

probability (i.e., incorrect option), would now be associated with a $p_R(B)=0.9$ probability (i.e., better option). The criterion remained the same as prior reversals. After reaching the criterion for the first probabilistic reversal learning phase or if rats were unable to achieve criterion after 10 days, the rats were advanced to the second probabilistic reversal phase (i.e., reversal 4, R4) beginning on the next testing day, where the probabilities would be reversed once again. As above for action-based reversal learning, rats that had previously received VEH during the first probabilistic reversal now received CNO injections, and vice versa.

Ex-vivo calcium imaging

In N=4 animals (all males), following >3 weeks of stereotaxic viral injections, rats (n=1 rat/brain region/virus combination; n=2-5 slices/rat) were deeply anesthetized with isoflurane (Patterson Veterinary, MA, USA), decapitated and brains submerged in ice cold bubbling slicing artificial cerebrospinal fluid (ACSF) containing (in mM): 62 NaCl, 3.5 KCl, 1.25 NaH_2PO_4, 62 choline chloride, 0.5 $CaCl_2$, 3.5 $MgCl_2$, 26 $NaHCO_3$, 5 N-acetyl L-cysteine, and 5 glucose, pH adjusted to 7.3 with KOH. Acutely microdissected vlOFC or BLA slices (300 μM thick) were obtained (VT1200s, Leica, Buffalo Grove, IL) and transferred to room temperature normal ACSF containing (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH_2PO_4, 2 $CaCl_2$, 2 $MgCl_2$, 26 $NaHCO_3$, 10 glucose, pH adjusted to 7.3 with KOH, and allowed equilibrate for >1 hr prior to the slice experiment. Slices were then transferred to chamber for imaging.

Imaging was performed on a Scientifica SliceScope, with imaging components built on an Olympus BX51 upright fluorescence microscope equipped with an sCMOS camera (Hamamatsu Orca Flash 4.0v3). Anatomical regions in brain sections for Ca^{2+} imaging were first identified by brightfield imaging with 780nm LED (Scientifica) illumination. Ca^{2+} imaging was performed using a 40x, 0.80NA water immersion objective (Olympus), continuous 470nm LED illumination

(ThorLabs), and a filter cube suitable for Calbryte 520AM imaging: Excitation: Brightline 466/40, Dichroic: Semrock FF495-Di03, Emission: Brightline 525/50. Images were acquired continually with 20ms exposure time. Electric field stimulation was applied at 110 mV (twin pulse every 5s). Temperature of ACSF during the recorded sessions was held at 28° C to minimize bubble formation.

Calcium data extraction

Prior to imaging sessions, 40x images of red and green fluorescence were captured and subsequently overlayed for post-hoc genotyping of individual cells ($GCaMP6f^{+}$, $GCaMP6f^{+}$/hM4Di-$mCherry^{+}$, or GCaMP6f/$mCherry^{+}$). Blinded scorers semi-manually curated regions of interest (ROIs) using Python-based Suite2P software (Pachitariu et al., 2017). ROI fluorescence was subtracted from the annual surround fluorescence, low-pass filtered, and transformed to dF/F_0 as previously described (Asrican & Song, 2021) where F_0 is calculated with a boxcar filter with a 200-frame lookback window. dF/F_0 values were clipped between 0 and 9000 to eliminate negative changes. Area under the curve and event frequency of each cell was calculated for each drug treatment. A threshold of 0.15 dF/F was used to determine significant events, which is lower than the dF/F of a single *ex-vivo* action potential, but significantly above signal to noise in our recorded traces (Tada et al., 2014), **Figure 4-1 C-F**.

Data Analyses

MATLAB (MathWorks, Natick, Massachusetts; Version R2021a) was used for all statistical analyses and figure preparation. Data were analyzed with a series of mixed-effects General Linear Models (GLM) for the discrimination learning phase to establish there were no

baseline differences in learning measures between the hM4Di and eGFP animals (i.e., virus group) prior to any drug treatment for each task separately. Mixed-effects GLMs were also conducted on reversal phases, with all fixed factors included in the model [i.e., reversal number (1-4), virus group (hM4Di, eGFP), drug (CNO, VEH), sex (female, male), drug order (CNO1, VEH1)] and individual rat as a random factor. These GLMs were run for each task type (stimulus- and action-based tasks) separately. Since learning reached asymptote at 5-days for stimulus-based reversal learning, only the first 5 days were included in the GLM. Similarly, since rats typically reached a plateau (and criterion) at 150 trials for action-based reversal learning, we included only the first 150 trials in the GLM. Significant reversal number and/or drug order interactions were further analyzed with a narrower set of fixed factors and Bonferroni-corrected post-hoc comparisons. In the instance where sex was found a significant predictor (moderator), sex was entered as a covariate factor in subsequent reversals. Accuracy (probability correct) before and after a reversal (-3 +3 sessions surrounding a reversal) was analyzed using ANOVA with virus group (hM4Di, eGFP) and drug order (CNO1, VEH1) as fixed factors on the average change pre-post reversal. Virus expression level was analyzed with ANOVA by sex (male, female) and region (vlOFC, BLA) on pixel counts.

Dependent measures for learning included probability of choosing the correct or better option, initiation latencies and omissions (failure to initiate a trial, latency to initiate a trial, respectively), correct and incorrect choice latencies (latency to select the correct or better stimulus or spatial side and latency to select the incorrect or wrong stimulus or spatial side, respectively), reward latencies (latency to collect the reward), probability of win-stay, and probability of lose-shift. Probability of win-stay and lose-shift adaptive strategies were calculated for the stimulus-based task such that each trial was classified as a *win* if an animal received a sucrose pellet, and as

a *loss* if no reward was delivered. Statistical significance was noted when p-values were less than 0.05. All Bonferroni post-hoc tests were corrected for number of comparisons.

To analyze *ex-vivo* calcium imaging data, 2-way ANOVAs with drug and virus as factors were conducted to compare calcium event changes in GCaMP6f and GCaMP6f+mCherry in each brain region for control experiments, and for GCaMP6f and GCaMP6f+hM4Di in each brain region for the experimental group. Tests corrected for number of comparisons were conducted for interactions.

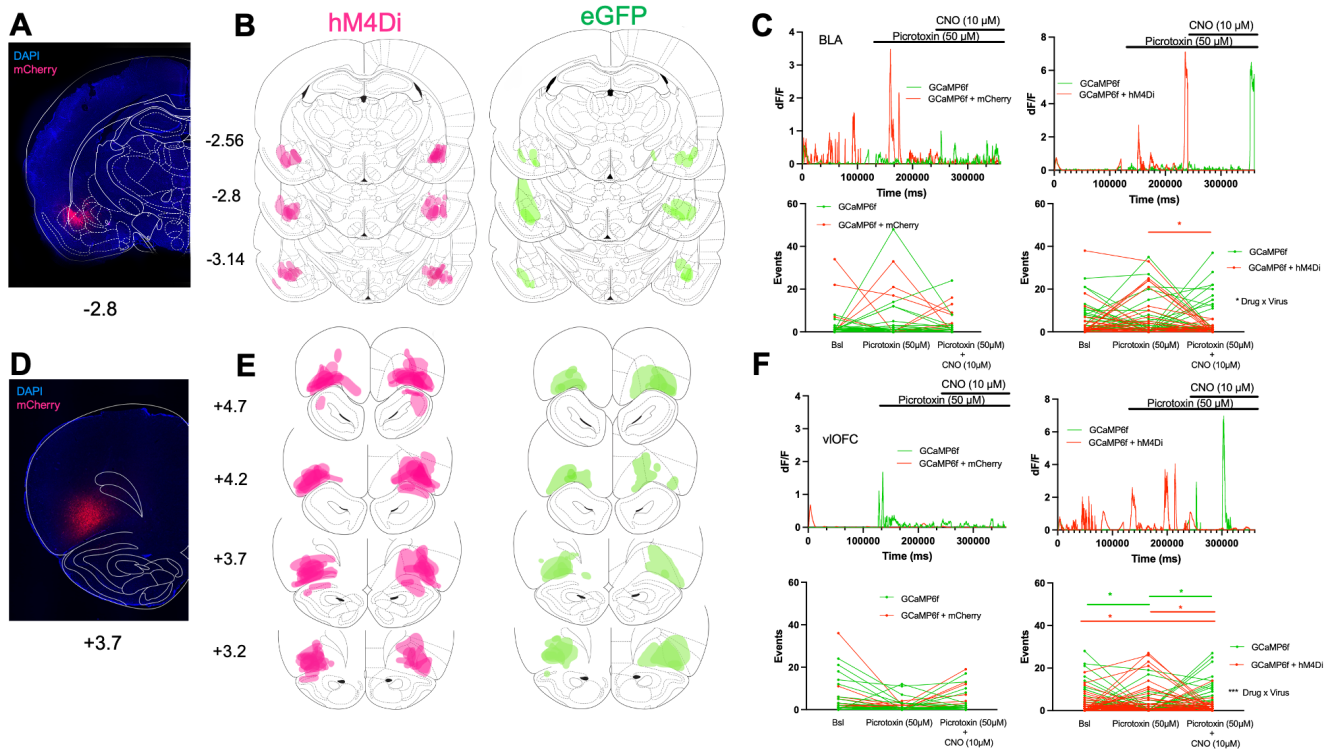

Figure 4-1. Bilateral targeting of basolateral amygdala and ventrolateral orbitofrontal cortex and confirmation of effective DREADDs inhibition via ex-vivo Ca^{2+} imaging in slices. (A) Photomicrograph of hM4Di-mCherry DREADDs expression in BLA. Numerals indicate mm anterior to Bregma. **(B)** Reconstructions of viral expression of hM4Di (magenta) and enhanced green fluorescent protein, eGFP (green) in BLA. The more intense colors represent regions where expression overlapped the most across animals. **(C)** In BLA neurons expressing GCaMP6f and GCaMP6f+mCherry, application of CNO (10μM) in the presence of picrotoxin (50 μM) had no effect on the frequency of elicited Ca^{2+} events (top left and right: example traces, bottom left: Ca^{2+} event changes). In BLA neurons that expressed hM4Di, there was a reduction in the frequency of elicited Ca^{2+} events during CNO application (bottom right). **(D)** Photomicrograph of hM4Di-mCherry DREADDs expression in vlOFC. Numerals indicate mm anterior to Bregma. **(E)** Reconstruction of viral expression of hM4Di (magenta) and enhanced green fluorescent protein, eGFP (green). The more intense colors represent regions where expression overlapped the most across animals. **(F)** In vlOFC neurons expressing GCaMP6f and GCaMP6f+mCherry, application of CNO (10μM) had no effect on the frequency of elicited Ca^{2+} events (top left and right: example traces, bottom left: Ca^{2+} event changes). In vlOFC neurons that expressed hM4Di, there was a reduction in the frequency of Ca^{2+} events during CNO application (bottom right). n=2-5 slices/rat, 2-way ANOVA and multiple comparison tests ***p<0.001, *p<0.05.

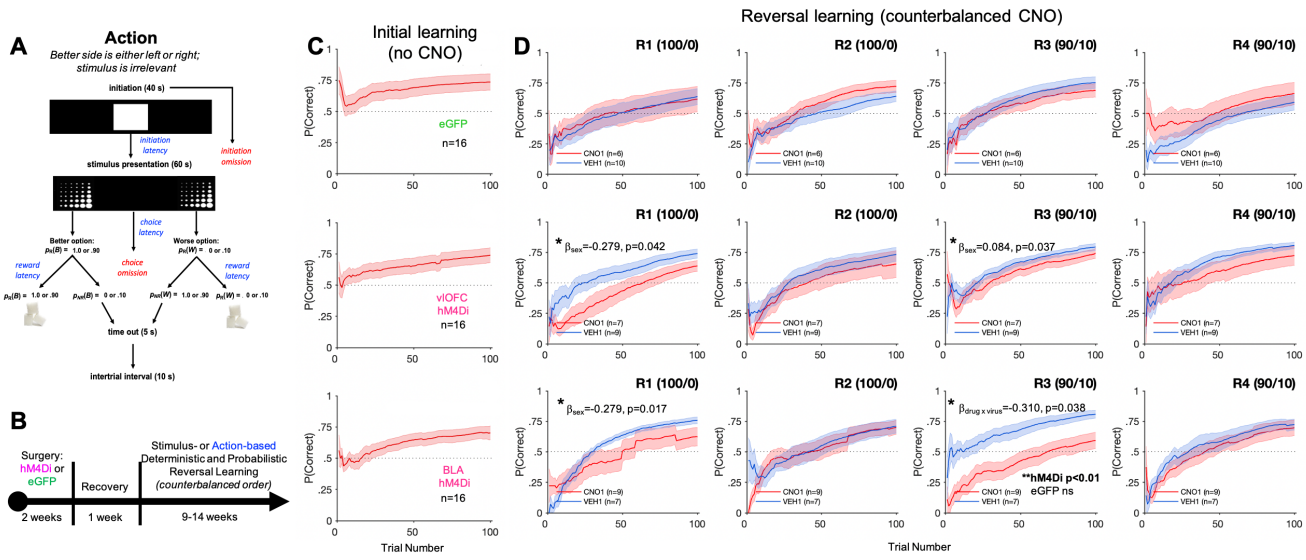

Figure 4-2. Either BLA or vlOFC inhibition attenuates action-based reversal learning. (A-B) Trial structure (A) and timeline (B) of the action-based task. Rats were first surgerized with either hM4Di DREADDs on a CaMKII promoter or eGFP null virus, also on the same promoter. Rats were allowed to recover for 1 week before testing on a stimulus-or action-based reversal learning task. **(C)** Initial learning of a rewarded side. **(D)** Learning during subsequent deterministic (100/0) and probabilistic (90/10) reversals. Plots show cumulative P(Correct) for first 100 trials with a sliding window of 10 trials is shown. Drug order was counterbalanced such that on R2 and R4 animals received VEH if they were administered CNO first on R1 and R3, and vice versa. There was no effect of CNO on learning in the eGFP group. *p<0.05 sex as a significant predictor (see **Fig. 3** for learning curves plotted by sex). Bonferroni-corrected post-hocs following mixed-effects GLM with sex as a covariate fixed factor wherein a drug x virus interaction was found resulted in **p<0.01 effect of drug only in BLA hM4Di, not in eGFP.

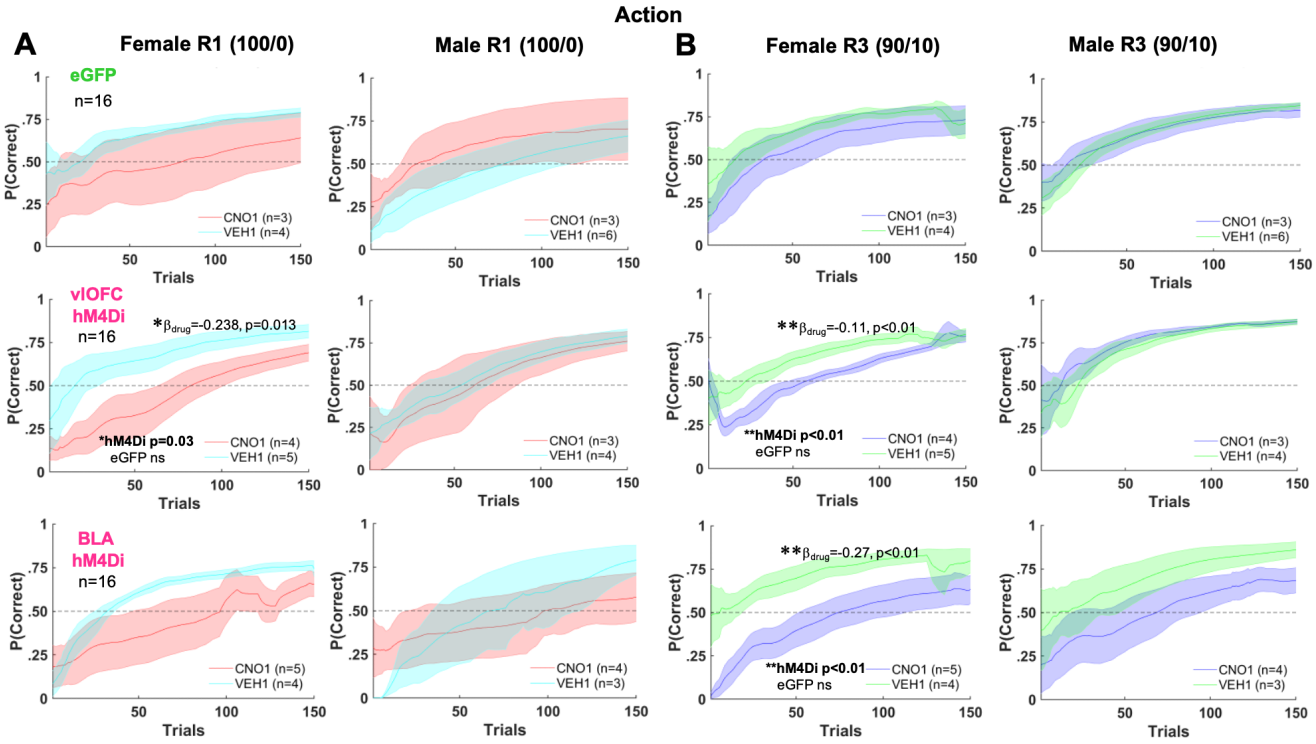

Figure 4-3. Female learning was more adversely affected by vlOFC and BLA inhibition than male learning during deterministic and probabilistic action-based reversals. Cumulative P(Correct) for first 150 trials with a sliding window of 10 trials is shown. Learning of deterministic (100/0) (A) and probabilistic (90/10) (B) reversals as measured by probability correct (P(Correct)). Drug order was counterbalanced such that on R2 and R4 (not shown) animals received VEH if they were administered CNO first on R1 and R3, and vice versa. Chemogenetic inhibition of vlOFC lowered P(Correct) on first deterministic R1 and first probabilistic reversal R3, whereas BLA inhibition attenuated probabilistic learning. Bonferroni-corrected post-hocs following GLM following a sex x drug interaction resulted in effect of drug only in vlOFC or BLA hM4Di, not in eGFP. There were no significant sex differences and no effect of CNO on learning in the eGFP group. **p<0.01, *p<0.05.

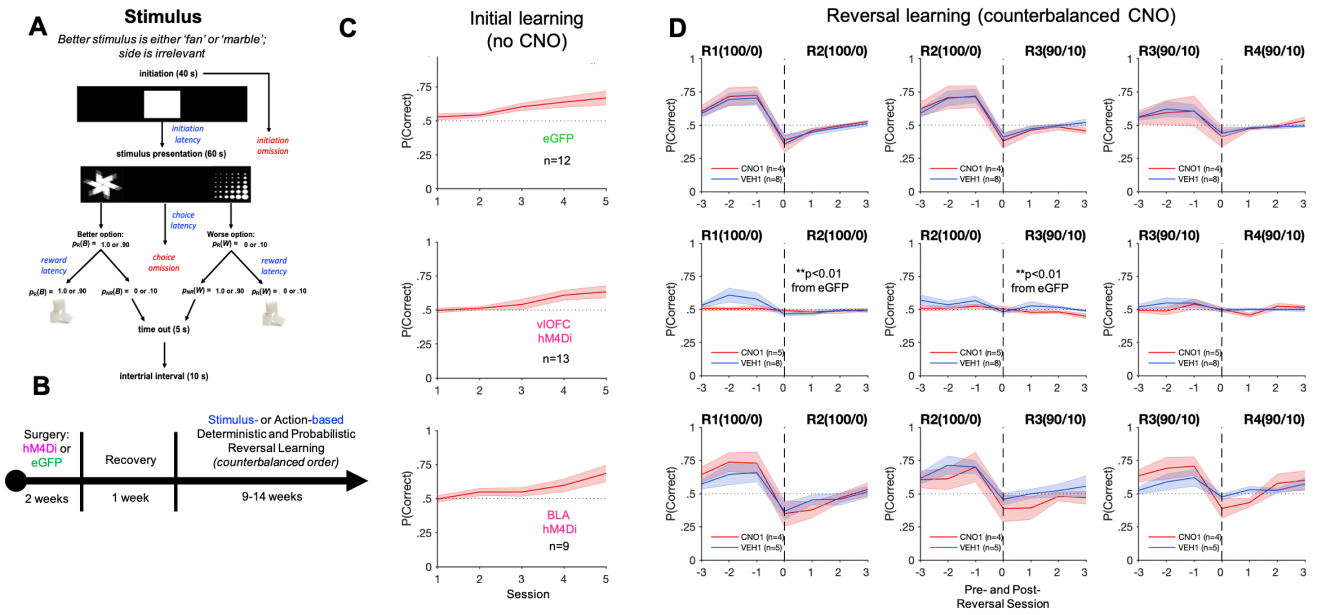

Figure 4-4. vlOFC, but not BLA, inhibition impairs the detection of stimulus-based reversals as measured by probability correct adjustments. **(A-B)** Trial structure (A) and timeline (B) of the stimulus-based task. Rats were first surgerized with either hM4Di DREADDs or eGFP null virus on a CaMKII promoter. Rats were allowed to recover for 1 week before testing on a stimulus-or action-based reversal learning task. **(C)** Initial learning of a rewarded stimulus, presented pseudorandomly on the left or right side of the touchscreen for eGFP (top), vlOFC hM4Di (middle), and BLA hM4Di (bottom). **(D)** Rats were always tested on a deterministic schedule before a probabilistic one. Shown are subsequent deterministic (100/0) and probabilistic (90/10) reversal "transitions," 3 sessions before and after each reversal. Drug order was counterbalanced such that on R2 and R4 animals received VEH if they were administered CNO first on R1 and R3, and vice versa. Chemogenetic inhibition of vlOFC abolishes changes in probability correct over the last 3 (pre) and first 3 (post)-reversal sessions, indicating impaired detection of reversal. In contrast, BLA inhibition had no impact on this detection. There was also no effect of CNO in eGFP group learning. **p<0.01 different than eGFP following ANOVA of pre-post difference.

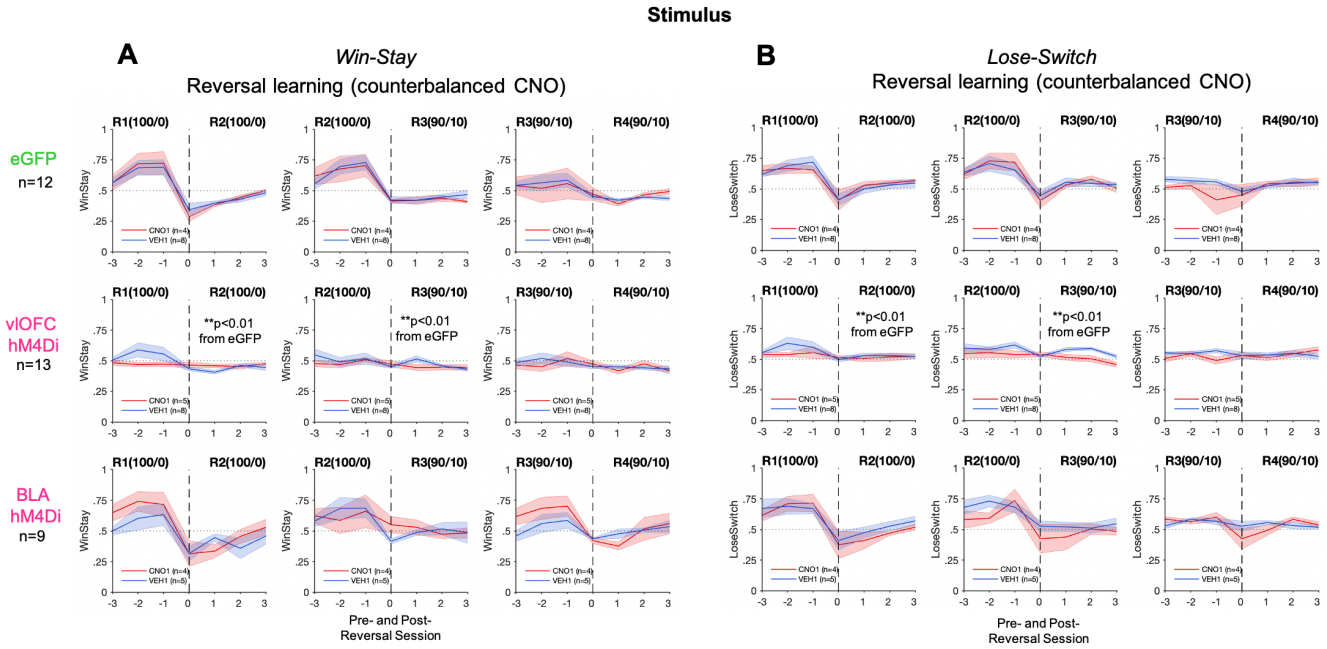

Figure 4-5. vlOFC, but not BLA, inhibition impairs the detection of deterministic stimulus-based reversals as measured by changes in Win-Stay, Lose-Switch strategies. (A) Inhibition of vlOFC abolishes changes in Win-Stay over the last 3 (pre) and first 3 (post)-reversal sessions, indicating impaired detection of R1-R2, R2-R3. In contrast, BLA inhibition had no impact on this detection. **(B)** Inhibition of vlOFC abolishes changes in Lose-Shift over the last 3 (pre) and first 3 (post)-reversal sessions, indicating impaired detection of R1-R2, R2-R3. In contrast, BLA inhibition had no impact on this detection. There was also no effect of CNO in eGFP group learning. **p<0.01 different than eGFP following ANOVA of pre-post difference.

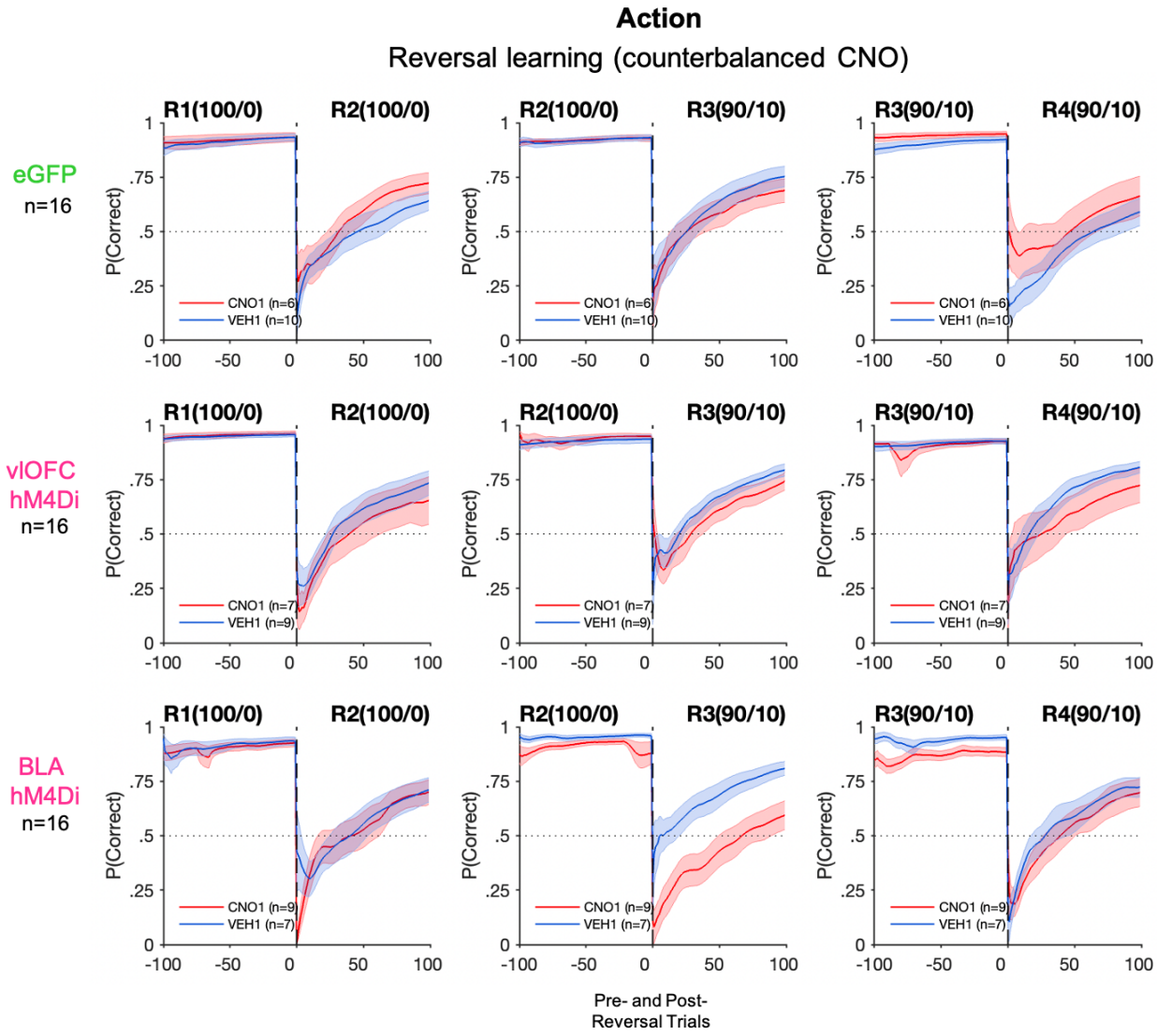

Figure 4-6. Neither BLA or vlOFC inhibition affected the detection of action-based reversals as measured by probability correct adjustments. Rats were always tested on a deterministic schedule before a probabilistic one. Shown are deterministic (100/0) and probabilistic (90/10) reversal "transitions," 100 trials before and after each reversal. Drug order was counterbalanced such that on R2 and R4 animals received VEH if they were administered CNO first on R1 and R3, and vice versa. Inhibition of BLA impaired probabilistic reversal learning as indicated by selectively poor performance in 90/10 at the beginning and end of that reversal. OFC inhibition resulted in no impairment around these reversals. There was also no effect of CNO in the eGFP group.

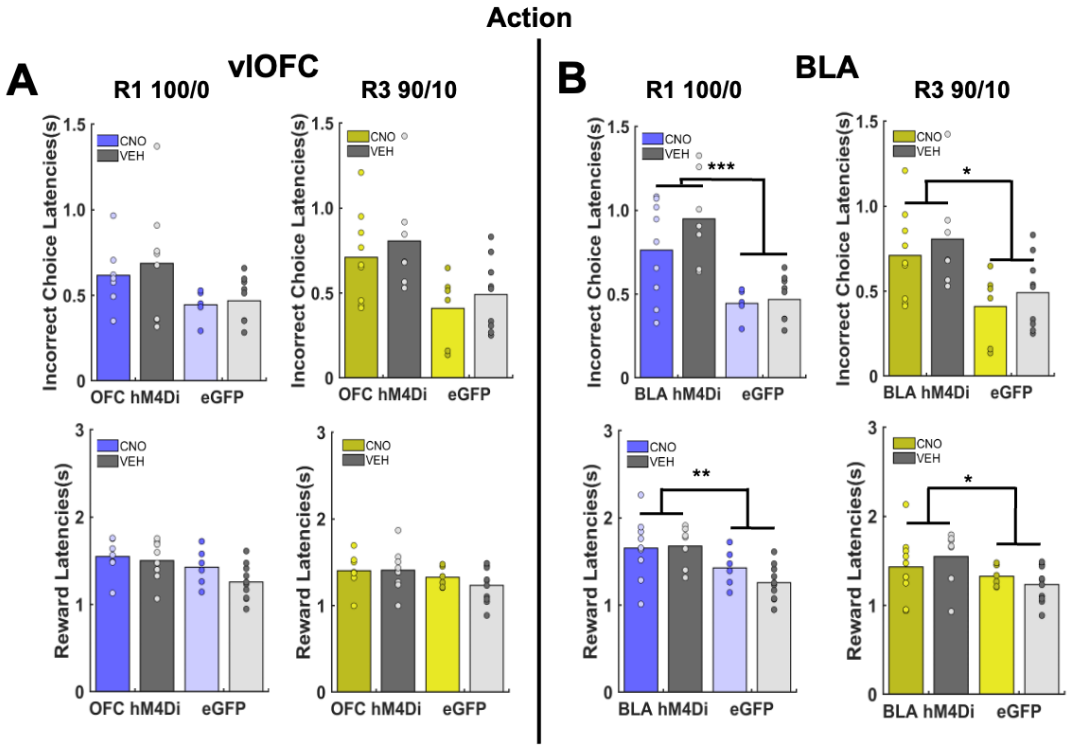

Figure 4-7. BLA, but not vlOFC, hM4Di exhibited longer incorrect choice and reward latencies than eGFP during action-based reversals. (A) There were no significant differences in incorrect choice and reward latencies when comparing vlOFC hM4Di vs. eGFP controls for any reversals. **(B)** BLA hM4Di animals exhibited longer incorrect choice and reward latencies than eGFP controls during reversals. ***p<0.0001, **p<0.001, *p<0.05

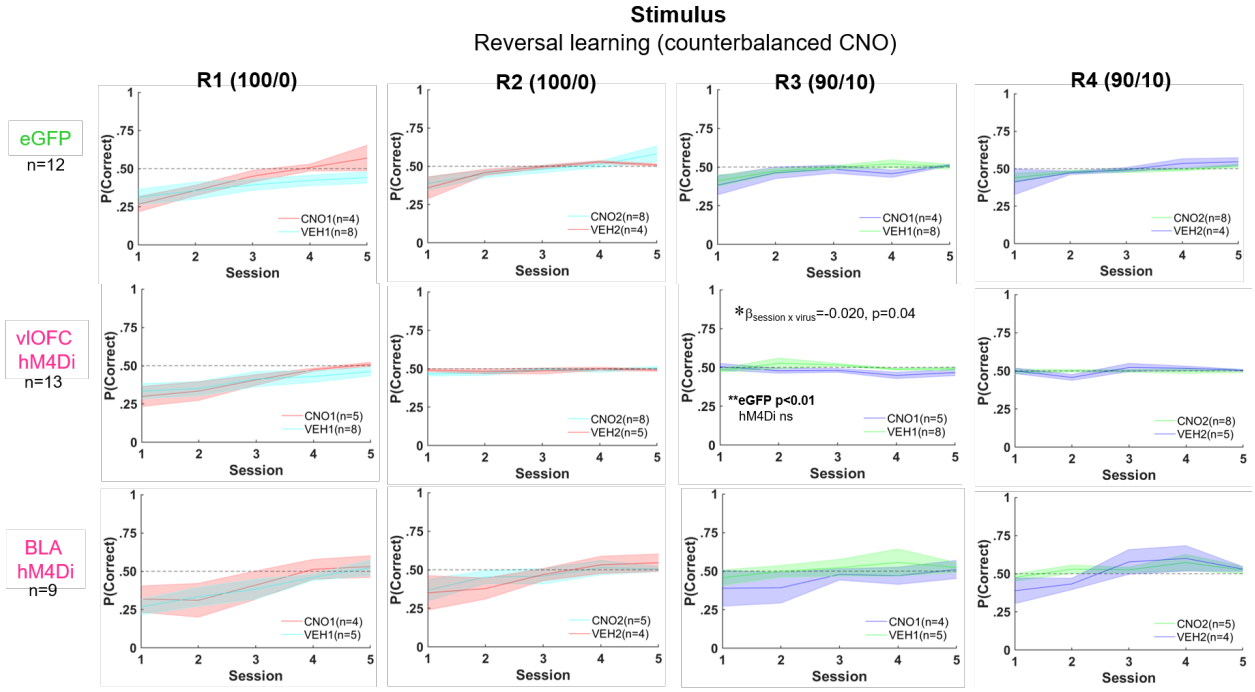

Figure 4-8. Probability correct during stimulus-based reversals. Cumulative P(correct) for first 5 sessions of each deterministic (100/0) and probabilistic (90/10) reversal. Drug order was counterbalanced such that on R2 and R4 animals received VEH if they were administered CNO first on R1 and R3, and vice versa. Despite most animals reaching criterion on initial learning (Fig. 2C), animals exhibited poor reversal learning. Performance around reversals (3 sessions before and after each reversal) is shown in Fig.2.

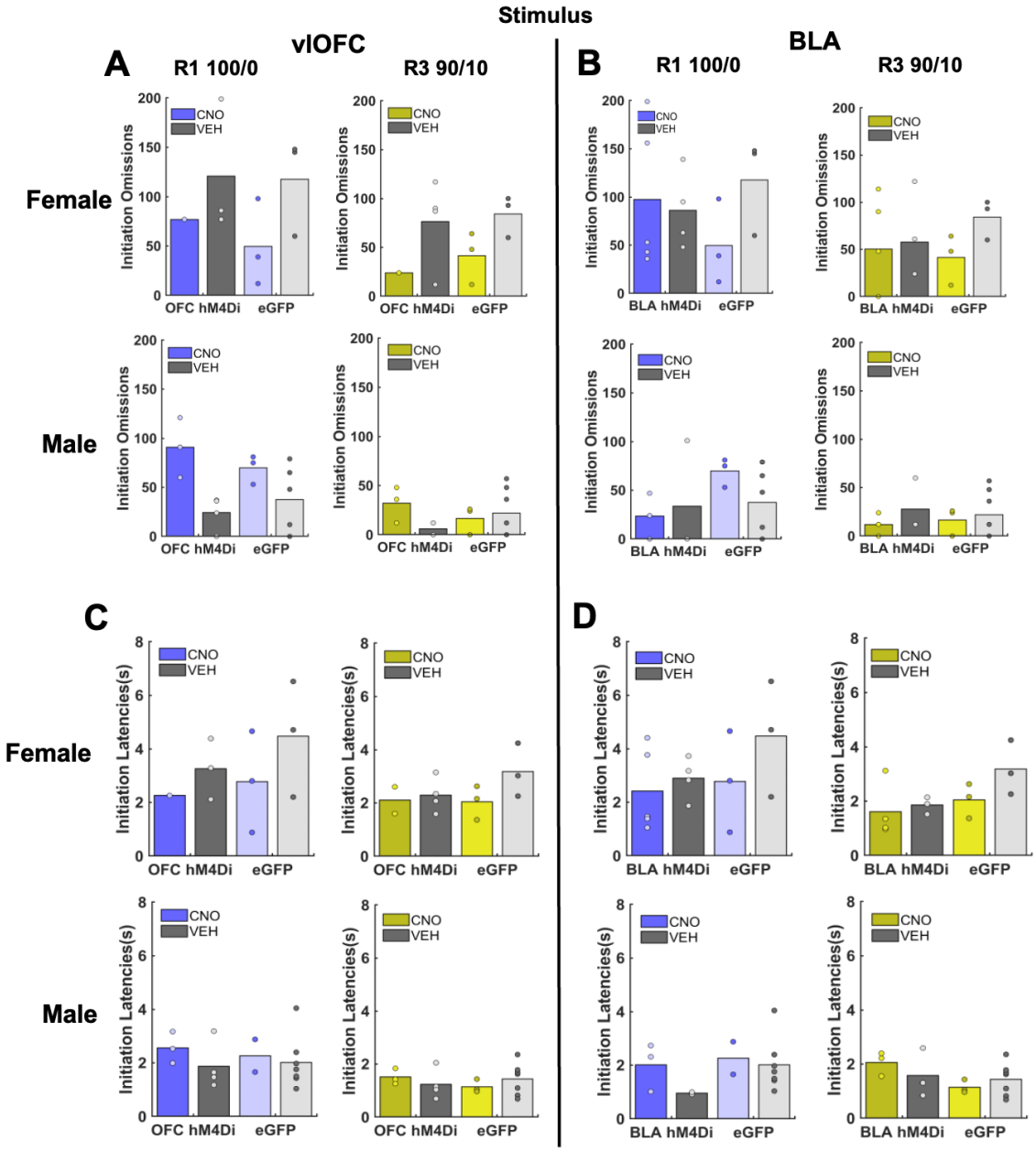

Figure 4-9. Females commit more initiation omissions and take longer to initiate trials than males in stimulus-based reversals. (A-B) Females committed more initiation omissions than males irrespective of virus group or drug condition across reversals. **(C-D)** Females take longer to initiate trials than males irrespective of virus group or drug condition across reversals.

Chapter 5: General Discussion

This work has investigated the behavioral and neural basis of flexible learning under uncertainty using a variety of approaches, including various manipulations of the reward environment (i.e., stimulus- vs action-based, deterministic vs. probabilistic reward outcomes) and techniques (i.e., voluntary alcohol administration, chemogenetic inhibition). In Chapter 2, I used a voluntary chronic intermittent model of alcohol administration to test the effects of prior alcohol exposure on flexible stimulus-based learning under probabilistic conditions (70/30). I found that females exhibited heavier alcohol consumptions patterns than males, and that alcohol-exposed animals were slower to learn, exhibited attentional deficits, and were less sensitive to negative feedback, compared to water-matched controls.

In Chapter 3, I compared the performance of rats on stimulus-based discrimination and reversal learning under two probabilistic conditions (70/30 and 90/30). I found all rats employed a win-stay adaptive strategy, during early discrimination, resulting in faster initial learning. Conversely, rat exhibited increased preservation during early reversal learning, resulting in slower learning. Additionally, a sex-dependent effect emerged on latencies to choose the better option, with females exhibiting longer latencies than males during reversal learning, an indication of slower processing speed. Finally, in collaboration with the Soltani Lab, we used a computational approach to fit choice behavior using reinforcement learning models and parameters, and found greater exploration and higher learning rates during reversal learning compared to initial discrimination learning.

In Chapter 4, I used a chemogenetic approach to inhibit pyramidal neurons in the ventrolateral orbitofrontal cortex (vlOFC) and basolateral amygdala (BLA), on both a stimulus- and action-based reversal paradigm under deterministic and probabilistic conditions. I found

mounting evidence of sex-modulated learning flexibility, such that females with either vlOFC or BLA inhibition were slower to learn during action-based reversals compared to males. Overall, the BLA seemed to be more necessary for learning action-based probabilistic reversals, as inhibition of this region resulted in poorer learning; whereas, the OFC was more necessary for detection of stimulus-based reversals.

Altogether, these findings highlight the importance of including sex as a biological variable, as all of these studies found sex-dependent effects that I will elaborate and expand upon below.

Sex differences in alcohol consumption patterns

I used the two-bottle choice procedure for the experiment in Chapter 1, as it is a reliable and efficient method of alcohol administration in animals, which promotes voluntarily ethanol consumption that may yield clinically relevant ethanol consumption patterns and dependence (Carnicella et al., 2014; Griffin, 2014; Nieto et al., 2021). Unlike continuous ethanol access which has limited external validity, placing limitations on ethanol availability using an intermittent schedule is thought to better mimic drinking patterns in humans, who often go extended periods of time without access to alcohol. In fact, there is a substantial evidence demonstrating that intermittent access to ethanol enhances intake in rodents across a variety of methods (Crabbe et al., 2011; Kimbrough et al., 2017), which corroborates the validity of intermittent administration in producing escalation of drinking, reflective of typical human alcohol consumption patterns.

Our findings from this study further validated prior research that has found sex-dependent difference in alcohol consumption, with females being more vulnerable to the effects of alcohol,

resulting in greater consumption and escalation over time (Hwa et al., 2011; Lourdes de la Torre et al., 2015; C. Vetter-O'Hagen et al., 2009; C. S. Vetter-O'Hagen & Spear, 2011). Some groups have reported that the rewarding effects of alcohol are enhanced in females compared to males, and suggest that increases in alcohol intake over time may be hormone-dependent (Torres et al., 2014). Erol and colleagues (2019) conducted a systematic review on the role of sex hormones on alcohol consumption and highlighted that the evidence in both the human and animal literature is rather inconclusive, with some groups reporting that estrogen levels either have no effect on or decrease alcohol intake (Cailhol & Mormede 2001; Ford, Eldridge, & Samson 2002b; Vetter-O'Hagen & Spear 2011), while others have found a positive relationships between estrogen levels and alcohol intake (Reid et al. 2002; Ford et al. 2002a, 2004; Marinelli et al. 2003; Reid et al. 2003; Quirarte et al. 2007; Rajasingh et al. 2007; Sherrill et al. 2011; Torres et al. 2014). Whereas, studies looking at the effects of testosterone and progesterone on alcohol consumption are even more scarce, and the few that have been conducted have found no association (Mankes et al., 1991; Almeida et al., 1998; Sinnott et al., 2002). In females, the association between menstrual cycle phase and alcohol consumption remains unclear, and although many report no association, the accuracy of results is questionable due to the sparsity of studies, small sample sizes and methodological inconsistencies in tracking menstrual cycles in humans (Warren et al., 2021; Carroll et al., 2015), with almost no studies conducted in rodents. The only rodent study to my knowledge that directly investigated the association between ethanol consumption and estrous cycle phase in rodents found that alcohol intake did not vary across the estrous cycle (Satta et al., 2018). Thus, more studies need to be conducted in order to increase the validity and generalizability of these results.

Sex differences in flexible stimulus-based learning with and without alcohol experience

Following early abstinence, females who had prior alcohol exposure, exhibited both learning and attentional deficits, as they required more pretraining sessions to meet criterion and committed more initiation omissions during initial discrimination learning than their male counterparts with the same experience. These alcohol-induced learning deficits did not extend to reversal learning, following prolonged abstinence from alcohol. Interestingly, irrespective of alcohol-experience, females committed more initiation omissions and exhibited longer initiation latencies than males, during late reversal learning.

Females exhibited greater attentional deficits during reversal learning than males, regardless of whether they had any prior alcohol experience. This suggests that the likely alcohol-induced sex-dependent deficits in learning and attention observed during early in learning (i.e., pretraining and discrimination) following acute abstinence, were transient and recoverable after longer abstinence periods, since reversal learning occurred at least a month after the last day of alcohol access. This finding is encouraging as it demonstrates that some alcohol-induced cognitive deficits may in fact be reversible and not as long-lasting; however, it is important to consider how the duration, amount, and frequency of drinking, as well as type of deficit observed, may modulate the severity of the deficit. It is also possible that withdrawal-induced stress, which in our study could have occurred during pretraining and possibly early discrimination learning, could have contributed to cognitive deficits, which then recovered after prolonged abstinence when withdrawal symptoms had subsided (Koob, 2008). For example, prior studies have found deficits in spatial memory after 24 hours (Pamplona-Santos et al., 2019), but not after 25 days of alcohol abstinence (Vetreno et al., 2020), in a Morris water maze task. Conversely, deficits in an object recognition task persisted for over 100 days following

abstinence, and were likely due to a longer alcohol exposure period (Vetreno and Crews, 2015). Therefore, it is important to consider the length of abstinence, as well as exposure period, when interpreting the extent and reversibility of cognitive deficits, as this can lead to variable results.

Sex differences in attention have not been largely supported by the literature, in fact, Grissom and Reyes (2019) conducted an extensive review of the literature on sex differences on several domains of executive function, including attention and did not find sufficient evidence to support sex differences in this particular domain. Although there is not substantial evidence to support sex differences in attention, there is evidence to support a potential EtOH-specific effect on attentional processing (Sanchez-Roige et al., 2014a; Irimia et al., 2014; Givens, 1997). Given the extensive literature examining the effects of ethanol exposure on attention using the 5-choice serial reaction time task (5-CSRTT), considered to be the golden standard for measuring attention in rodents, it seems plausible ETOH-exposed animals could be experiencing attention deficits, impairing their ability to initiate trials in a timely matter. A 2014 study evaluating attentional capacity using the 5-CSRT55 following ETOH exposure, found ETOH-exposed rats had more omissions than ETOH-naïve rats, while motivation remained intact, with no differences in correct responses or reward latency, similar to our findings (Irimia et al., 2014), however, it is important to note these differences were observed only during acute, not prolonged abstinence from ETOH exposure. Other studies have found an ETOH dose-dependent decrease in choice accuracy, specifically impairments in ability to direct and sustain attention to brief stimuli, but not a complete disruption on overall performance (i.e., percentage correct), suggesting an impairment in attentional processing (Givens, 1997), which further corroborate our findings. Therefore, it is plausible that sex effects we observe here are moderated by EtOH-

experience, resulting in more pronounced deficits in attentional processing in EtOH-drinking females.

Sex differences on performance measures is not modulated by reward environment

Performance measures assessing differences in latencies, employment of adaptive strategies, and perseverative behavior revealed sex-dependent effects. During reversal learning, females exhibited longer choice latencies, and were more likely to employ the win-stay adaptive strategy than males, even when they had previously chosen the worse, less rewarded option. Females also exhibited more perseverative behavior than males, as indicated by higher repetition index scores, during discrimination learning; this pattern was also observed during reversal learning, although it was only a marginal effect. No sex-dependent effects were observed in reinforcement learning models.

Although correct response or choice latencies are not typically reported in most studies, they are often used as a measure of processing or decision-making speed. Our findings are consistent with a recent study that found that males had shorter prepotent responses latencies than females in an action-based reversal learning task, although this was only tested under deterministic and not probabilistic conditions, as was done in our study (Aarde et al., 2020). Orsini et al. (2017) did an extensive review on sex differences in animal models of decision-making, and found that in tasks closely model uncertainty (i.e., an unknown probability distribution), male rodents seem to be using more global information to make decision, while females assess both options first by switching between advantageous and disadvantageous options before determining the most adaptive choice (van den Bos et al. 2012; Orisini et al., 2017). In tasks that closely model risk (i.e., known probability distribution), females developed an optimal choice more quickly than males

(Peak et al., 2015; Orsini et al., 2017). Thus, female choice behavior may initially take longer to determine the optimal choice, but once it is established, their choice latency is likely to decrease compared to males.

Sex differences in use of employment of adaptive strategies like win-stay and lose-shift has not been extensively studied, although recent papers from the Grissom lab have found that female mice displayed more win-stay behaviors than males, as they were more likely to choose the same option that was rewarded on a previous trial on a spatial-based two-armed bandit probabilistic task (Chen et al., 2021a), which is consistent with our study. Females were also more likely to use spatial-based win-stay and lose-shift strategies, but eventually switched to using these adaptive strategies based on the visual stimuli on a two-armed visual bandit task (Chen et al., 2021b). And finally, greater perseveration during early reversal is consistent with the long-standing idea that reversal learning requires inhibitory control, such that the inability to disengage from previously rewarding behavior after a change in contingency may be reflective of impulsive response tendencies (Izquierdo & Jentsch, 2012). Indeed, there is evidence that inflexible responding in reversal learning may be genetically related to impulsivity (Crews & Boettiger, 2009; Fineberg et al., 2010; Franken et al., 2008; Groman et al., 2009; Groman & Jentsch, 2013; Jentsch et al., 2014). However, studies looking at the effects of sex on perseverative behavior in reversal learning have been mixed with some finding that female rats exhibit more perseverative behavior (Gargiulo et al., 2022), as was seen in our study, while other have found that male rodents actually commit more perseverative errors during reversal learning (Aarde et al. 2021). Overall, males and females seem to use different strategies during reversal learning, while sex-dependent effects on perseverative behavior remain inconclusive.

Females exhibited poorer action-based learning following vlOFC inhibition

In our study, we found that OFC inhibition resulted in learning impairments during action-based deterministic and probabilistic reversals, but in a sex-dependent manner. Several studies have found that inactivation and lesions of the OFC impair action-based reversal learning by decreasing the total number of reversals achieved, as rats require more sessions to reach criterion, and the use of adaptive strategies, like win-stay and lose-shift (Verharen et al., 2020; Dalton et al., 2016), as well as increasing the animals' incorrect responses, perseverative errors (Boulougouris et al., 2007), and choice latencies (Dalton et al., 2016) under both deterministic and probabilistic conditions. Unfortunately, we are unable to compare our sex-dependent findings on reversal learning following OFC inhibition given that all the aforementioned studies were conducted in male rodents. However, there is one study done on marmosets investigating sex differences on a stimulus-based reversal learning task, that found females required more sessions to reach criterion during reversal learning than males, and exhibited greater perseverative behavior (LaClair et al., 2019), but it is unclear if these findings would extend to action-based reversal learning, as was used in our study. Additionally, some studies in humans have investigated whether OFC activity is modulated by fluctuations in ovarian hormones throughout the menstrual cycle, and found greater OFC activity during the anticipation of uncertain rewards (Dreher et al., 2007), and certain (but delayed) reward (Bayer et al., 2013) in the follicular phase. Altogether, these findings demonstrates that sex-dependent modulation of brain activity during learning can occur, but whether this extends to reversal learning and/or rodents is unknown due to the paucity of literature on sex differences.

Females exhibited poorer action-based learning following BLA inhibition

In our study, we found that BLA inhibition resulted in learning impairments only during probabilistic action-based reversals in a sex-dependent manner. Although studies on sex-dependent effects of BLA activity on decision-making have been even more sparse, there have been several studies that have found sex differences on dopaminergic and corticotropin releasing factor receptors (Williams et al., 2021; Salvatore et al., 2018), that may elucidate some of the neuronal mechanisms underlying our results. Females have higher mRNA expression of both dopamine receptor 2 (D2R) and corticotropin releasing factor receptor 1 (CRFR1) in the amygdala than males, with only CRFR1 expression was negatively correlated with advantageous responding in females (Georgiou et al., 2018). In a probabilistic reversal learning task, increases in central CRF decreased motivation by increasing trial omissions and decreasing choice latencies, as well as reducing sensitivity to negative feedback following a reversal, this effect was more robust in females than males (Bryce and Floreco, 2021). Although we did not find sex differences in performance measures like omission and latencies, it is still important to consider the extent to which these measures affect learning accuracy over time, and how this may be affected by hormones, like CRF.

Sex differences in dopaminergic systems have also been found in rodents, with male having a higher density of dopaminergic synaptic boutons within the BLA than females, although the density of dopaminergic axons did not differ by sex (Manion et al., 2022). In reversal learning, lower levels of D2R expression were associated with impairments in rodents (DeSteno and Schmauss, 2009), and the effect of D2R stimulation on reversal learning was modulated by orbitofrontal-amygdalar-striatal networks in humans (van der Schaaf et al., 2013). However, because these studies only used male subjects, it is unknown whether there is similar dopamine receptor expression and neuronal networks associated with reversal learning in females. Sex

hormones can also act as neuromodulators, affecting the levels of dopamine in females (Barth et al., 2015). Estradiol has been shown to enhance dopamine synthesis and release, while decreasing reuptake, resulting in higher levels of dopamine (Hwang et al., 2020; Del Rio et al., 2018; Yoest et al., 2014, 2019; Becker, 2004; Jacobs et al. 2011). Thus, it is plausible amygdala activation may be modulated by fluctuations in ovarian hormones that occur throughout the menstrual cycle, as it has some of the highest densities of estrogen- and progesterone-receptors in the brain (Barth et al., 2015), which may impact dopaminergic systems. In fact, greater amygdala activation in anticipation of uncertain reward occurred during the follicular phase (Dreher et al., 2007), whereas amygdala activation after receiving a large monetary reward was positively correlated with changes in ovarian hormones (Macoveanu et al., 2016). Overall, these findings highlight the importance of considering sex differences on dopaminergic systems and associated neural networks, as well as how these systems may be modulated by hormones, like CRF and estrogen. These dynamic interactions may affect the neuronal activity of regions recruited during reversal learning, and thus can have important implications in terms of the interpretation and generalizability of findings.

And finally, the learning impairment observed in females following both vlOFC and BLA inhibition may not be primarily due to fluctuations in hormone levels in females relative to males, but rather may actually reflect differences in strategies between sexes (Grissom and Reyes, 2019; Chen et al., 2021a, 2021b). Several studies from the Grissom lab found that although males and females achieved the same level of performance in terms of accuracy, in both action(spatial)-based (Chen et al., 2021a) and stimulus-based (Chen et al., 2021b, Cox et al, 2023) probabilistic reversal learning paradigms, there were different sex-dependent strategies that emerged. These studies found that female rodents learned quicker during exploration, as they would exploit the better

option much earlier in learning than males even though they explored less overall than males (Chen et al. 2021a, 2021b). Conversely, males explored more and were more inconsistent (i.e., switching their choice frequently), and because they explored for longer periods of time, they exploited the better option much later in learning compared to females (Chen et al., 2021a, 2021b). Cox et al. (2023) also found that males and females learned at similar rates in an auditory stimulus-based probabilistic reversal learning task, but they differed in initiation latencies, such that females displayed less motivation to engage in the task for lower value trials, a difference which we did not observe in our study. The emergence of sex-dependent effects in the employment of strategies during reversal learning further merits the need to use both male and female in all studies.

Females exhibit attentional deficits in stimulus-based reversals relative to males

Sex did not moderate performance in learning in stimulus-based reversals as it did in action-based reversals; however, females did committed more initiation omissions and latencies across all reversals (i.e. except R2 and R4 in BLA hM4Di/eGFP comparisons) in all virus group regardless of inhibition, suggesting potential attentional deficits consistent with our own prior findings (Aguirre et al., 2020), perhaps brought about by the tendency for females to avoid negative outcomes (Orsini et al., 2016, 2017; van de Bos et al., 2012). Although recent studies have suggested that failure to initiate a trial may not only be a proxy for attention (Aguirre et al., 2020), but might also be considered a measure of motivation or task engagement (Cox et al., 2023; Wang et al., 2013; Hamid et al., 2016). These studies suggest that females were slower to engage in the task (i.e., longer initiation latencies) following an unrewarded trial or in lower value trials. Cox et al. (2023) also monitored the estrous cycle of female rodents to assess whether this modulated the differences in motivation between males and females and found that females were slower on trials

with a low relative chosen value during the follicular phase when estradiol and progesterone are high (Cox et al., 2023), consistent with other studies that found a similar effect (Verharen et al., 2019).

Conclusions and Future Directions

Altogether, these findings highlight the importance of including sex as a biological variable in any experimental design, as all of these studies found sex-dependent effects that provided insight on how drugs, like alcohol, impact male and females differently, as well as how differences in flexible learning, may reflect sex differences in the employment of learning strategies and recruitment of associated brain regions.

After extensively reviewing the literature on sex differences in order to contextualize our own study findings, it became very apparent that the majority studies only included males across species, thus, making it hard to interpret and generalize our findings. Although the preference of using male subjects over females was not surprising, as historically male animals have been six times more likely to be used than females (Beery et al, 2011; Will et al., 2017), it was still surprising that some of the more recent literature also did not include females. Hence, large gaps in our knowledge and understanding of these sex differences remain, due to the sparsity of studies done with female subjects. Given our own sex-dependent findings in the context of flexible learning, it is possible that the findings and interpretations of previous studies that have only used male animals in similar studies, could be vastly different in females, and thus, may not be as generalizable as previously thought.

There are several sex-specific effects that have been found in some studies, but whose validity and generalizability is unclear due to the paucity of studies, and thus, require further

probing and replication. Firstly, neuronal plasticity which has been strongly implicated in learning, and can be sex-specific depending on experience, for example, some studies have found that chronic stress exposure affecting dendritic and spine morphology in both prefrontal cortex and amygdala (Farrell et al., 2015, Shanksy et al., 2009). Future studies should consider the possibility of sex-dependent changes in neuronal plasticity and how this may impact learning. Secondly, although the majority of studies have not found sex differences in learning accuracy, there have been sex differences on how information is gathered and the type of strategies used (Chen et al., 2021a, 2021b; Cox et al., 2023). Specifically, males prefer to sample more in order to build a representation of the rules in the environment, while females prefer to learn a single dimension at a time and exploit the most rewarding option sooner (Chen et al., 2021a). Future studies should consider using computational modeling, including reinforcement learning models, that can identify and predict the types of learning strategies used by both males and females. Thirdly, the majority of studies have found females tend to be more sensitive to negative feedback (Cox et al., 2023; Orsini et al., 2022; Pellman et al., 2017; Van den Bos et al., 2012). However, it remains unclear, whether sex modulates neuronal activity associated with feedback sensitivity. For example, inhibition of lOFC reduces learning from both positive and negative feedback in males (Verharen et al., 2020), while BLA-lesioned male rats show feedback insensitivity following positive and negative outcomes, and thus are less likely to use adaptive strategies (Stolyarova et al., 2017). Future studies should investigate whether these findings extend to females as well. Fourthly, the frontal cortex and amygdala are regions known to be impacted by the effects of alcohol at different stages of the addiction cycle (Koob 2014). The frontal cortex has been shown to be compromised by alcohol, resulting in deficits in executive function, including behavioral flexibility, and is associated with enhanced craving in response to drug-related cues over naturally rewarding

reinforcers (Koob and Le Moal, 1997; Koob, 2014; Koob and Volkow, 2016); whereas, the amygdala responds to drug-conditioned cues, and is thought to be heavily involved in cue-induced drug reinstatement (Koob and Le Moal, 1997; Koob, 2014; Koob and Volkow, 2016). Given that these regions are both involved in flexible learning, future studies should investigate how alcohol disrupts and/or alters neuronal activity in these brain regions, and how these changes may contribute to subsequent cognitive effects following alcohol experience.

It is important to consider the effects of the menstrual cycle in females on learning and decision-making. Unfortunately, the potential of estrous-driven behavioral variation in females has commonly been cited as the rationale for excluding female rodents in behavioral neuroscience research (Shansky et al., 2021; Beery et al., 2011; McCarthy et al., 2012), however, the majority of studies have not found this to be the case in risk-taking behavior (Islas-Preciado et al., 2020; Wallin-Miller et al., 2017) or exploratory behavior (Lovick et al., 2021; Meziane et al., 2007). Future studies should investigate whether estrous cycle modulates other types of behavior, as this is still not fully understood. It is also important to consider key differences in the menstrual cycle in primates and estrous cycle in rodents, as they may impact the translatability of such studies, this includes, the shorter average length of estrous cycle in rodents compared to the menstrual cycle in primates, and lack of the luteal phase in rodents (Finn, 2020). Finally, although there are several methods available used to track estrous cycle in rodents, they are not frequently implemented, as many are time-consuming, laborious, and/or unreliable (Ajayi and Akhigbe, 2020). Vaginal cytology is considered the gold standard in the field as it is the most commonly used method to track estrous cycle, but other methods, like urine biochemistry and vaginal wall impedance, require improvements to increase the feasibility and accuracy of these techniques (Ajayi and Akhigbe, 2020). Altogether, these findings should encourage the inclusion of both sexes, especially given

several meta-analyses have found that females do not exhibit greater behavioral variation than males (Becker et al., 2016; Prendergast et al., 2014; Beery, 2018; Kaluve et al., 2022). Additionally, improving the quality, accuracy, efficiency, and reliability of techniques used to track menstrual cycle phase, would increase the number of studies investigating the role of hormones on alcohol consumption patterns and flexible learning. This in turn, would provide further insight into the underlying mechanisms behind the sex-dependent effects that have emerged in several studies, including ours.

The National Health Institute's (NIH) policy to include sex as a biological variable (SABV) in all NIH-funded research was implemented in 2016 (NIH, 2016), and this has only recently begun to improve (Mamlouk et al, 2020; Woitowich et al., 2020). This policy has led to a real shift in the scientific community, motivating many to re-evaluate how own inherent biases, which may have resulted in flawed experimental design, and affected the interpretation and generalizability of our findings. The emergence of sex-dependent studies in the few studies that have included both sexes, the NIH's SABV policy implemented in 2016, and the literature demonstrating that females do not have greater behavioral variability than males, should ameliorate any remaining concerns in the neuroscience research community and encourage the inclusion of both sexes. These collective efforts will result in a more comprehensive approach to conducting neuroscience research, as future scientific discoveries could benefit the health of both males and females, leading to better science, and ultimately moving the field forward.

APPENDIX A: Chapter 3 Computational Modeling and Analyses

RESULTS

Comparisons of Repetition measures

Comparisons between early Discrimination & Reversal Learning. We next fitted GLMs to examine the effects of diferent factors on repetition in choice behavior. We did not find any significant effect of phase, group, sex, or any significant interaction of those factors on overall $p(Stay)$, p(Stay | better), or p($Stay \mid worse$) (**Appendix B Table 12**). However, there was an increased perseveration index in the reversal phase relative to discrimination (**Figures A-1A and A-1E**), regardless of reward probability group or sex (**Appendix B Table 13**). For *RI,* we found a significantly lower value in the 70-30 group and in males, but no significant effects of phase, phase*group, phase*sex, group*sex, or phase*group*sex interaction (**Figures A-1B and A-1F**). Similarly, for RI_B (**Appendix B Table 14**) we observed lower values in the 70-30 group and in males, but no significant difference by phase, phase*group, phase*sex, group*sex, or phase*group*sex interaction (**Figures A-1C and A-1G).** This pattern holds for RI_W as well, with a lower RI_W in the 70-30 group and males, and no significant effects of phase, phase*group, phase*sex, group*sex, or group*phase*sex interaction (**Figures A-1D and A-1H**). As we observed no phase interactions, we were not justified in further analyses of discrimination and reversal phases, separately.

Summary. We found more perseveration during early reversal learning, compared to the discrimination phase. This indicates animals continued to perform according to contingencies learned in discrimination. When using the RI measure, we observed less repetition (lower RI, particularly RI_W) in the 70-30 reward probability group and in males, suggesting probability group

and sex have an effect on repetitive behavior that are not due to differences in the propensity to be rewarded.

Comparisons of Estimated Parameters based on Fit of Choice Behavior with RL models

Comparisons between Discrimination & Reversal Learning. To gain more insight into learning and decision making, we next estimated model parameters (learning rate, α and sensitivity to difference in subjective reward values, σ) based on two RL models across groups in each of the two phases. Unlike our analyses of response to reward feedback and repetition measures, we include all sessions in our RL analysis rather than the first seven to capture learning over a longer time period. We found that the single-learning rate model, RL1, fit the choice data significantly better, as shown by the lower BIC value (difference in BIC between RL1 and RL2 = -7.78, pairwise t-test: $t(49) = -75.8$, $p < 0.0001$). For this reason, only results for RL1 are presented below.

Comparison of estimated parameters from RL1 revealed that the average learning rate was significantly higher during reversal than discrimination phase (difference in α =0.28; two sample t-test: $t(48)=-5.48$, $p<0.0001$) and at the same time, sensitivity to subjective reward values was significantly lower (difference in σ=-0.77; two sample t-test: $t(48)=8.18$, $p<0.00001$), which corresponds to enhanced exploration. Additionally, to better explore the relationship between learning rate and sensitivity to difference in reward values, we calculated the correlation coefficients between parameters across groups. We found a significant negative correlation between learning rate and sensitivity to difference in reward values during reversal (r=-0.76, p=0.0028) but not discrimination learning (r=-0.42, p=0.0762). We also calculated correlation between model parameters using the inverse Hessian at the ML estimate (see Methods). However, we did not find any evidence that the correlations between model parameters of the RL1 model to be significantly different from 0 during discrimination (r=-0.16±0.20, $t(24) = -0.82$, $p = 0.42$) or

reversal learning (r=0.017±0.14, t(24) = 0.12, p = 0.90). These results suggest that the observed simultaneous increase in learning rates and decrease in sensitivity to subjective reward value in the reversal relative to discrimination phase was not due to our fitting procedure and, instead, happened due to independent mechanisms.

Comparisons within Discrimination Learning. Using estimated parameters based on RL1, we found that the 90-30 reward probability group had a learning rate $\alpha = 0.038 \pm 0.017$, sensitivity to difference in reward values $\sigma = 0.92 \pm 0.12$ and fit of $BIC = 2227 \pm 600$ (**Figure A-2**). The 70-30 reward probability group had a learning rate $\alpha = 0.032 \pm 0.031$, sensitivity to difference in reward values $\sigma = 0.86 \pm 0.12$ and fit of $BIC = 4234 \pm 622$. We found no significant differences between reward probability groups in terms of α (two sample t-test: $t(23)$= 0.74, p=0.47) or σ (two sample t-test; $t(23)$= -0.32; p=0.76).

Comparisons within Reversal Learning. As in the discrimination phase, we again fit choice behavior for each subject in the reversal phase using RL1 (**Figure A-2B**). We found that the 90-30 reward probability group had a learning rate $\alpha = 0.36 \pm 0.069$, sensitivity to difference in reward values $\sigma = 0.082 \pm 0.056$ and fit of $BIC = 3167 \pm 390$. The 70-30 reward probability group had a learning rate $\alpha = 0.26 \pm 0.061$, sensitivity to difference in reward values $\sigma = 0.13 \pm 0.04$ and fit of $BIC = 5262 \pm 396$. We found no significant difference between the α (two sample t-test; $t(23)$= -0.96; p=0.35) or σ (two sample t-test; $t(23)$ = 0.70; p=0.49) parameters between groups.

Summary. In analyses across discrimination and reversal phases, we found that sensitivity to difference in reward values was lower during reversal than discrimination, and learning rate(s) were higher (the single learning rate for RL1, and both learning rates for RL2) during reversal relative to the discrimination phase. These results indicate that reversal caused faster learning and more exploration at the same time. However, within each phase of learning, the estimated model

parameters were not significantly different between probability groups in either phase. We also found no sex-dependent differences. This suggests the slight increased probability of reward corresponding to the better option in the 90-30 reward probability group, as compared to the 70-30 group, was not large enough to induce significantly different learning or, alternatively, this effect could not be captured by our RL models.

METHODS

Data Analyses

Metrics of repetition and perseveration. We also used two additional higher-order measures: repetition index and perseveration index. We calculated repetition index (Soltani et al., 2013) as the difference between the actual probability of staying and the chance level of staying:

$$RI = P(stay) - (P(better) \times P(better) + P(worse) \times P(worse))$$

where $P(stay)$ is the actual probability of staying equal to the joint probability of choosing the option on two consecutive trials, $P(better)$ is the probability of choosing the more rewarded stimulus, and $P(worse)$ is the probability of choosing the less rewarded stimulus. We further parsed repetition index into RI_B and RI_W, which accounts for differing tendency to stay on better and worse options:

$$RI_B = P(stay, better) - (P(better) \times P(better))$$

$$RI_W = P(stay, worse) - (P(worse) \times P(worse))$$

Additionally, we introduced a perseverative index, analogous to the perseverative index defined in Brigman et al. (2008) and Brigman, Mathur, et al. (2010) as the number of same stimulus choices following a loss divided by the number of such "runs," or the ratio of first-presentation stimulus errors to consecutive errors.

We used these two measures, in addition to the probability of staying, to analyze repetitive behavior. As above, we only used the first seven sessions in calculating repetition measures.

Reinforcement learning model. We utilized two simple reinforcement learning models to capture animals' learning and choice behavior across all sessions during discrimination and reversal. In each model, the estimated subjective reward value of a choice (*V*) was determined by prior choices and corresponding reward feedback. Specifically, subjective reward values were updated based on reward prediction error, the discrepancy between actual and expected reward. For each observed choice, we updated the subjective reward value of the corresponding option. Accordingly, if the more rewarded stimulus was chosen then $V \ = V_{better}$, and if the less rewarded stimulus was selected $V \ = V_{worse}$, where V was the option subjective reward value. We used the following learning rules to update V_{Choice}:

RL1: Model with a single learning rate. On a given trial t, the subjective reward value of the chosen stimulus is updated using the following function:

$$V(t+1) = V(t) + \alpha\big(r(t) - V(t)\big)$$

Where $V \ (t)$ is the subjective reward value on trial t, $r(t)$ is 1 (0) if trial t is rewarded (not rewarded) and α is the single learning rate.

RL2: Model with separate learning rates for rewarded and unrewarded trials. On a trial t, the following learning rule was used:

$$V(t+1) = \begin{cases} V(t) + \alpha_{Rew}\big(1 - V(t)\big), r(t) = 1 \\ V(t) - \alpha_{Unrew} V(t)), \quad r(t) = 0 \end{cases}$$

where α_{Rew} and α_{Unrew} are the learning rates for rewarded and unrewarded trials, respectively.

We then applied the following decision rule to determine the probability of selecting the better option on trial t:

$$P_{better}(t) = \frac{1}{1+exp(-\sigma(V_{better}-V_{worse}))},$$

where σ is the inverse temperature parameter or sensitivity to the difference in subjective reward values.

We fit each model's parameters to minimize the negative log likelihood of observed responses. The learning rate parameters (α, α_{Rew}, and α_{Unrew}) were restricted to values between 0 and 1, and σ was constrained to values greater than 0. Learning rate parameters were passed through a sigmoid function to avoid local minima at very low learning rates. Initial parameter values were selected from this range, then fit using MATLAB's *fmincon* function. For each set of trials, we performed 20 iterations with different initial, randomly selected parameter values to avoid local minima, and the best fit was selected from the iteration with the lowest negative log likelihood (*LL*). Additionally, we also calculated the Bayesian information criterion (BIC) for each fit, defined as

$$BIC = kln(\mathrm{n}) - 2(LL)$$

Where *k* is the number of parameters (two in RL1, and three in RL2) and *n* is the number of trials.

We used two methods to examine the relationship between the learning rates and sensitivity to difference in subjective reward values. First, we calculated the Pearson correlation coefficient between estimated learning rate and inverse temperature across animals. That is, after estimating parameters for both groups during discrimination and reversal, we calculated the correlation between sets of learning rates and inverse temperatures across groups during each phase. Second, we examined the parameter correlation derived from our parameter fitting procedure. More specifically, for each animal in each phase, we used the inverse of the Hessian matrix output by *fmincon* to estimate parameter covariance. From this covariance matrix we calculated the

correlation matrix. This allowed us to obtain analytic estimates of the correlation between parameters for each animal in each phase (Bishop, 2006; Daw, 2011).

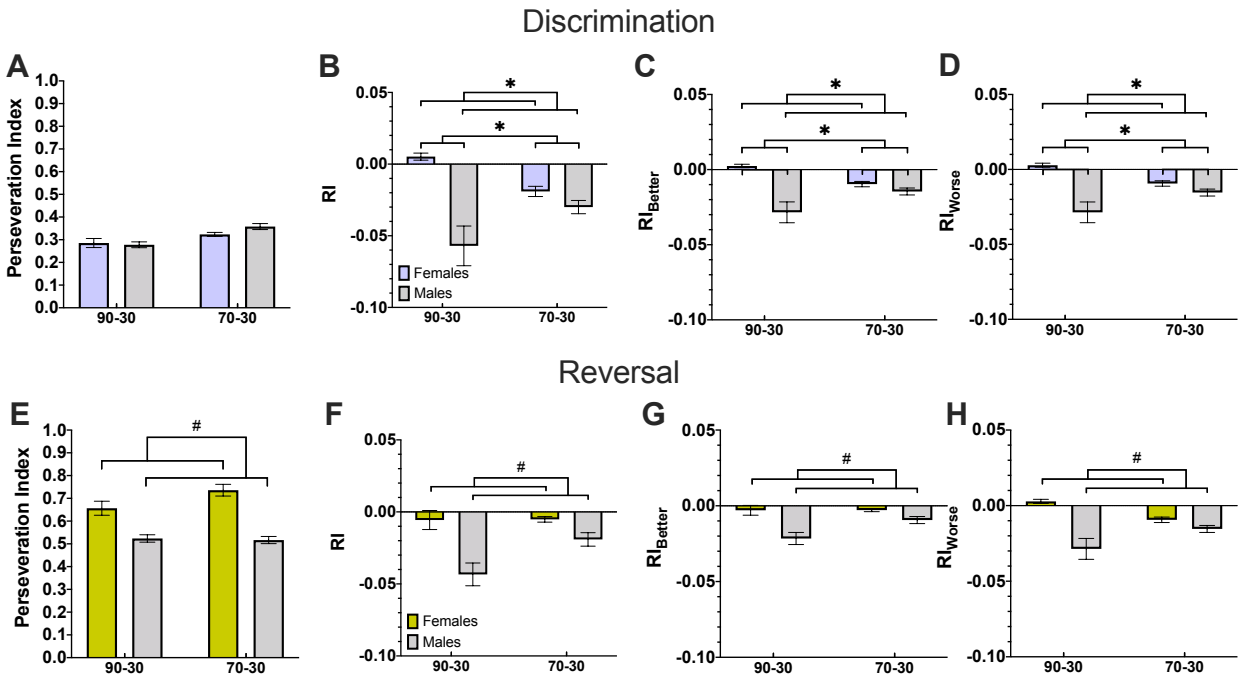

Figure A-1. Greater perseveration during reversal and lower repetition index measures for males as compared to females in both phases, and for the 70-30 group as compared to the 90-30 group in the discrimination phase. Plotted is the perseveration index (A), overall repetition index (B), repetition index for the better option (C), and repetition index for the worse option (D) in the discrimination phase. Though we find no significant differences in perseveration index, the 70-30 group shows a significantly lower RI and RI_B, and marginally significantly lower RI_W. Additionally, males have significantly lower values for all three repetition index measures. **(E-H)** Same as A-D, but for reversal. We observe a significantly lower perseveration index and a marginally significant lower RI, RI_B, and RI_W in males than females. Bars indicate $\pm$ SEM $^{\#}$p=0.06, *p $\leq$0.05

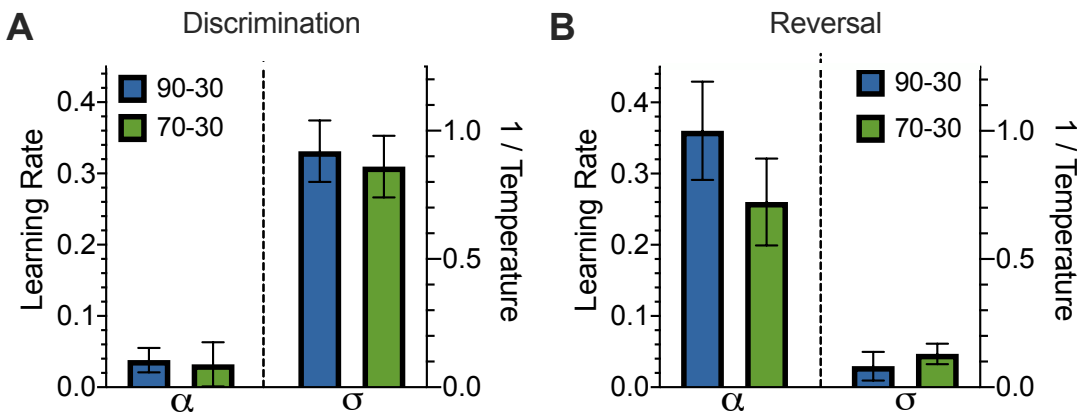

Figure A-2. Higher learning rate and lower sensitivity to difference in subjective reward values in reversal compared to discrimination learning. (A-B) Learning parameters and sensitivity to difference in reward values for the single-learning rate model during the discrimination (A) and reversal (B) phases. We find no significant difference in parameter values between reward probability groups during discrimination or reversal. However, we do find significantly higher learning rates and significantly lower sensitivity to difference in reward values parameters following reversal.

APPENDIX B: Chapter 3 Data Tables

Table 1. *Probability of choosing the better option during early discrimination and reversal learning*

Early Discrimination and Reversal Learning γ = probability of choosing the better option							
Formula	γ~[1+phase*group*sex*day+(1|rats)]						
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.4479	0.0463	9.6847	328	<.0001	0.3570	0.5389
phase	-0.2293	0.0758	-3.0271	328	**0.0027**	-0.3783	-0.0803
group	0.0070	0.0535	0.1314	328	0.8955	-0.0982	0.1122
sex	0.0433	0.0763	0.5667	328	0.5713	-0.1069	0.1934
day	0.0462	0.0104	4.4607	328	**<.0001**	0.0258	0.0666
sex:group	-0.0270	0.0848	-0.3189	328	0.7500	-0.1938	0.1398
sex:phase	0.0836	0.1000	0.8357	328	0.4039	-0.1132	0.2804
group:phase	-0.0270	0.0831	-0.3251	328	0.7453	-0.1905	0.1365
sex:day	-0.0270	0.0228	-1.1846	328	0.2370	-0.0719	0.0179
group:day	-0.0157	0.0142	-1.1069	328	0.2692	-0.0435	0.0122
phase:day	-0.0040	0.0157	-0.2532	328	0.8003	-0.0348	0.0269
sex:group:phase	0.0134	0.1138	0.1175	328	0.9065	-0.2105	0.2372
sex:group:day	0.0152	0.0254	0.5979	328	0.5503	-0.0348	0.0652
sex:phase:day	0.0086	0.0247	0.3474	328	0.7285	-0.0400	0.0572
group:phase:day	0.0070	0.0193	0.3611	328	0.7183	-0.0310	0.0449

Table 2. *Total number of rewards collected during early discrimination and reversal learning*

Early Discrimination and Reversal Learning γ =number of rewards							
Formula	γ~[1+phase*group*sex*day+(1|rats)]						
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	50.727	10.730	4.7278	328	<.0001	29.620	71.835
phase	-0.6417	11.562	-0.0555	328	0.9558	-23.387	22.103
group	-2.1959	12.206	-0.1799	328	0.8573	-26.207	21.815
sex	29.130	17.266	1.6871	328	0.0925	-4.8371	63.097
day	5.9808	2.0728	2.8854	328	**0.0042**	1.9031	10.058
sex:group	-27.929	21.903	-1.2751	328	0.2032	-71.018	15.160
sex:phase	-21.182	29.708	-0.7130	328	0.4764	-79.624	37.260
group:phase	-5.3005	15.352	-0.3453	328	0.7301	-35.502	24.901
sex:day	-2.4540	5.8143	-0.4221	328	0.6733	-13.892	8.9840
group:day	-2.1913	2.4551	-0.8925	328	0.3728	-7.0211	2.6386
phase:day	-1.4236	2.2873	-0.6224	328	0.5341	-5.9233	3.0760
sex:group:phase	23.830	33.182	0.7182	328	0.4732	-41.446	89.106
sex:group:day	-2.0186	6.094	-0.3312	328	0.7407	-14.007	9.9698
sex:phase:day	0.0086	0.0247	0.3474	328	0.7285	-0.0400	0.0572
group:phase:day	0.0070	0.0193	0.3611	328	0.7183	-0.0310	0.0449

Table 3. *Initiation and choice omissions during early discrimination and reversal learning*

Early Discrimination and Reversal Learning Formula: γ~[1+phase*group*sex+(1|rats)]							
γ =initiation omissions							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	138.40	40.087	3.4524	42	0.0013	57.500	219.30
phase	-18.600	17.622	-1.0555	42	0.2972	-54.162	16.962
group	-27.900	44.500	-0.6270	42	0.5341	-117.71	61.906
sex	-58.650	51.131	-1.147	42	0.2579	-161.84	44.538
phase:group	69.350	48.505	1.4298	42	0.1602	-28.537	167.24
phase:sex	1.1000	28.122	0.0391	42	0.9690	-55.652	57.852
group:sex	164.90	70.480	2.3397	42	**0.0241**	22.665	307.13
phase:group:sex	-158.10	61.395	-2.5751	42	**0.0136**	-282.00	-34.200
γ =choice omissions							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	1.000	0.2828	3.5355	42	0.0010	0.4292	1.5708
phase	0.200	0.3347	0.5976	42	0.5533	-0.4754	0.8754
group	0.625	0.8003	0.7809	42	0.4392	-0.9902	2.2402
sex	3.750	2.7390	1.3691	42	1.3691	-1.7774	9.2774
phase:group	-0.700	0.5751	-1.2172	42	0.2303	-1.8606	0.4606
phase:sex	-2.450	2.0085	-1.2198	42	0.2293	-6.5032	1.6032
group:sex	-4.625	2.8545	-1.6202	42	0.1127	-10.386	1.1357
phase:group:sex	2.575	2.0843	1.2354	42	0.2235	-1.6313	6.7813

Table 4. *Initiation and better choice latency during early discrimination and reversal learning*

Early Discrimination and Reversal Learning							
Formula: γ~[1+phase*group*sex+(1|rats)]							
γ=initiation latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	4.5118	0.8013	5.6307	42	<.0001	2.8947	6.1289
phase	-0.8195	0.4883	-1.6782	42	0.1007	-1.8049	0.1660
group	-0.2730	0.9112	-0.2996	42	0.7660	-2.1118	1.5658
sex	-1.1703	1.0509	-1.1136	42	0.2718	-3.2911	0.9505
phase:group	1.0286	0.9268	1.1098	42	0.2734	-0.8418	2.8991
phase:sex	-0.5760	0.9377	-0.6143	42	0.5423	-2.4683	1.3163
group:sex	3.3462	1.7256	1.9392	42	0.0592	-0.1362	6.8287
phase:group:sex	-2.4249	1.7569	-1.3802	42	0.1748	-5.9706	1.1207
γ=better choice latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.9025	0.0704	12.818	42	<.0001	0.7604	1.0446
phase	0.0332	0.0737	0.4503	42	0.6548	-0.1156	0.1820
group	0.1099	0.1540	0.7138	42	0.4793	-0.2009	0.4208
sex	-0.1401	0.1273	-1.1005	42	0.2774	-0.3971	0.1169
phase:group	-0.2615	0.1338	-1.9537	42	0.0574	-0.5315	0.0086
phase:sex	-0.2118	0.0785	-2.6991	42	**0.0100**	-0.3702	-0.0534
group:sex	-0.0796	0.2092	-0.3804	42	0.7056	-0.5017	0.3426
phase:group:sex	0.2999	0.1648	1.8193	42	0.0760	-0.0328	0.6325

Table 5. *Worse choice and reward latency during early discrimination and reversal learning*

Early Discrimination and Reversal Learning Formula: γ~[1+phase*group*sex+(1|rats)]							
γ =worse choice latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.8044	0.0824	9.7582	42	<.0001	0.6380	0.9708
phase	0.0965	0.0312	3.0977	42	**0.0035**	0.0336	0.1594
group	0.0885	0.1358	0.6519	42	0.5181	-0.1856	0.3627
sex	-0.1177	0.1553	-0.7575	42	0.4530	-0.4311	0.1958
phase:group	-0.1059	0.0919	-1.1521	42	0.2558	-0.2913	0.0796
phase:sex	-0.1543	0.0828	-1.8636	42	0.0694	-0.3213	0.0128
group:sex	-0.0308	0.2051	-0.1502	42	0.8814	-0.4446	0.3830
phase:group:sex	0.0965	0.1379	0.6999	42	0.4879	-0.1818	0.3748
γ =reward collection latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	1.9392	0.1232	15.737	42	<.0001	1.6905	2.1879
phase	-0.0482	0.0677	-0.7120	42	0.4804	-0.1848	0.0884
group	-0.5845	0.1387	-4.2155	42	**0.0001**	-0.8643	-0.3047
sex	-0.1688	0.2034	-0.8300	42	0.4113	-0.5793	0.2417
phase:group	0.0313	0.0781	0.4001	42	0.6911	-0.1264	0.1890
phase:sex	-0.2257	0.0906	-2.4909	42	**0.0168**	-0.4085	-0.0428
group:sex	0.2706	0.2202	1.2288	42	0.2260	-0.1738	0.7151
phase:group:sex	0.2924	0.1188	2.4604	42	**0.0181**	0.0526	0.5322

Table 6. *Initiation omissions, better choice latency, and reward collection latency during discrimination learning*

Discrimination Learning Formula: γ~[1+group*sex]							
γ =initiation omissions							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	236.00	145.72	1.6195	21	0.1203	-67.05	539.05
group	76.125	185.76	0.4098	21	0.6861	-310.19	462.44
sex	-11.000	218.59	-0.0503	21	0.9603	-465.58	443.58
group:sex	462.88	272.63	1.6978	21	0.1043	-104.08	1029.8
γ =better choice latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.9173	0.1052	8.7342	21	<.0001	0.6989	1.1357
group	0.0555	0.1339	0.4147	21	0.6826	-0.2229	0.3339
sex	-0.0619	0.1575	-0.3931	21	0.6982	-0.3895	0.2657
group:sex	-0.1631	0.1965	-0.8303	21	0.4157	-0.5717	0.2455
γ =reward collection latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	1.9705	0.1200	16.42	21	<.0001	1.7209	2.2201
group	-0.6006	0.1530	-3.9258	21	**0.0008**	-0.9187	-0.2824
sex	-0.2429	0.1800	-1.3492	21	0.1916	-0.6172	0.1315
group:sex	0.3960	0.2245	1.7638	21	0.0923	-0.0709	0.8629

Table 7. *Initiation omissions, better choice latency, and reward collection latency during reversal learning*

Reversal Learning Formula: γ~[1+group*sex]							
γ =initiation omissions							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	381.80	113.35	3.3682	21	0.0029	146.07	617.53
group	222.08	144.50	1.5369	21	0.1393	-78.426	522.58
sex	-189.30	170.03	-1.1133	21	0.2782	-542.90	164.30
group:sex	-167.08	212.07	-0.7878	21	0.4396	-608.09	273.94
γ =better choice latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.8899	0.0930	9.5727	21	<.0001	0.6966	1.0832
group	-0.1766	0.1185	-1.4901	21	0.1511	-0.4230	0.0699
sex	-0.3229	0.1394	-2.3156	21	**0.0308**	-0.6129	-0.0329
group:sex	0.2153	0.1739	1.2382	21	0.2293	-0.1463	0.5770
γ =reward collection latency							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
Intercept	1.8941	0.1442	13.139	21	<.0001	1.5943	2.1939
group	-0.4881	0.1838	-2.6560	21	**0.0148**	-0.8703	-0.1059
sex	-0.3879	0.2163	-1.7936	21	0.0873	-0.8376	0.0619
group:sex	0.4887	0.2697	1.8121	21	0.0843	-0.0722	1.0496

Table 8. *Initiation omissions, better choice latency, and reward collection latency during reversal learning with Discrimination Sessions to Criterion (dis_stc) as a covariate*

Reversal Learning (Covariate: Discrimination Sessions to Criterion)							
Formula: γ~[1+group*sex+dis_stc]							
γ = initiation omissions							
Coefficients	**β**	SE	tStat	DF	p	CIL	CIU
intercept	445.62	116.21	3.8344	20	0.0010	203.2	688.04
group	272.93	142.18	1.9197	20	0.0693	-23.643	569.50
sex	-125.48	167.93	-0.7472	20	0.4636	-475.77	224.81
dis_stc	-7.9773	5.2301	-1.5253	20	0.1428	-18.887	2.9325
group:sex	-123.20	204.87	-0.6014	20	0.5544	-550.55	304.15
γ = better choice latency							
Coefficients	**β**	SE	tStat	DF	p	CIL	CIU
intercept	0.8958	0.0996	8.9951	20	<.0001	0.6881	1.1036
group	-0.1719	0.1218	-1.4106	20	0.1737	-0.4260	0.0823
sex	-0.3170	0.1439	-2.2026	20	**0.0395**	-0.6172	-0.0168
dis_stc	-0.0007	0.0045	-0.1653	20	0.8570	-0.0101	0.0086
group:sex	0.2194	0.1756	1.2498	20	0.2258	0.1468	0.5856
γ = reward collection latency							
Coefficients	**β**	SE	tStat	DF	p	CIL	CIU
intercept	1.8448	0.1521	12.131	20	<.0001	1.5275	2.162
group	-0.5274	0.1861	-2.8348	20	**0.0102**	-0.9155	-0.1393
sex	-0.4372	0.2198	-1.9894	20	0.0605	-0.8956	0.0212
dis_stc	0.0062	0.0068	0.9009	20	0.3784	-0.0081	0.0204
group:sex	0.4548	0.2681	1.6965	20	0.1053	-0.1044	1.014

Table 9. *Win-stay and lose-shift during early discrimination and reversal learning*

Early Discrimination and Reversal Learning Formula: γ~[1+phase*group*sex+(1|rats)]							
γ =win-stay							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.5941	0.0122	48.508	42	<.0001	0.5694	0.6188
phase	-0.1269	0.0290	-4.3795	42	**0.0001**	-0.1853	-0.0684
group	-0.0746	0.0344	-2.1708	42	**0.0357**	-0.1440	-0.0052
sex	-0.0876	0.0551	-1.5900	42	0.1193	-0.1988	0.0236
phase:group	0.1068	0.0512	2.0834	42	**0.0433**	0.0033	0.2102
phase:sex	0.0442	0.0621	0.7118	42	0.4805	-0.0811	0.1695
group:sex	0.0403	0.0667	0.6036	42	0.5494	-0.0944	0.1750
phase:group:sex	-0.0328	0.0782	-0.4196	42	0.6769	-0.1907	0.1251
γ =lose-switch							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.5721	0.0156	36.586	42	<.0001	0.5405	0.6036
phase	-0.1572	0.0370	-4.2478	42	**0.0001**	-0.2318	-0.0825
group	-0.0634	0.0192	-3.3082	42	**0.0019**	-0.1020	-0.0247
sex	0.0398	0.0186	2.1392	42	**0.0383**	0.0023	0.0773
phase:group	0.0345	0.0487	0.7082	42	0.4827	-0.0638	0.1328
phase:sex	0.0325	0.0409	0.7942	42	0.4316	-0.0501	0.1150
group:sex	-0.0321	0.0253	-1.2661	42	0.2124	-0.0832	0.0191
phase:group:sex	0.0368	0.0561	0.6564	42	0.5152	-0.0763	0.1499

Table 10. *Win-stay, win-stay|better, and win-stay|worse during discrimination learning*

Discrimination Learning							
Formula: γ~[1+group*sex+(1|rats)]							
γ =Win-Stay							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.5941	0.0122	48.508	21	<.0001	0.5686	0.6196
group	-0.0746	0.0344	-2.1708	21	**0.0416**	-0.1461	-0.0031
sex	-0.0876	0.0551	-1.5900	21	0.1268	-0.2022	0.0270
group:sex	0.0403	0.0667	0.6036	21	0.5526	-0.0985	0.1791
γ =Win-Stay|better							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.6349	0.0145	43.817	21	<.0001	0.6047	0.6650
group	-0.0857	0.0412	-2.0801	21	0.0500	-0.1714	<.0001
sex	-0.0971	0.0619	-1.5674	21	0.1320	-0.2258	0.0317
group:sex	0.0288	0.0779	0.3699	21	0.7152	-0.1332	0.1908
γ =Win-Stay|worse							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.3791	0.0220	17.266	21	<.0001	0.3334	0.4247
group	0.0034	0.0360	0.0957	21	0.9247	-0.0715	0.0784
sex	-0.0402	0.0350	-1.1481	21	0.2638	-0.1131	0.0326
group:sex	0.0853	0.0574	1.4858	21	0.1522	-0.0341	0.2046

Table 11. *Win-stay,* win-stay|better, and win-stay|worse *during reversal learning*

Reversal Learning Formula: γ~[1+group*sex+(1|rats)]							
γ =Win-Stay							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.4672	0.0174	26.850	21	<.0001	0.4311	0.5034
group	0.0304	0.0256	1.1878	21	0.2482	-0.0228	0.0836
sex	-0.0434	0.0228	-1.9004	21	0.0712	-0.0909	0.0041
group:sex	0.0092	0.0333	0.2771	21	0.7844	-0.0600	0.0784
γ =Win-Stay|better							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.3822	0.0351	10.901	21	<.0001	0.3093	0.4552
group	-0.0696	0.0542	-1.2843	21	0.2130	-0.1822	0.0431
sex	0.0147	0.0392	0.3744	21	0.7119	-0.0668	0.0962
group:sex	0.0671	0.0597	1.1238	21	0.2738	-0.0571	0.1913
γ =Win-Stay|worse							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.6115	0.0232	26.313	21	<.0001	0.5632	0.6598
group	0.0468	0.0450	1.0390	21	0.3106	-0.0469	0.1404
sex	-0.1253	0.0285	-4.4024	21	**0.0002**	-0.1844	-0.0661
group:sex	0.0208	0.0593	0.3505	21	0.7294	-0.1026	0.1442

Table 12. *P(stay), p(stay|better), and p(stay|worse) during early discrimination and reversal learning*

Early Discrimination and Reversal Learning γ~[1+phase*group*sex+(1|rats)]							
γ =P(Stay)							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.5420	0.0054	100.37	42	<.0001	0.5311	0.5529
phase	-0.0172	0.0174	-0.9867	42	0.3295	-0.0523	0.0179
group	-0.0335	0.0234	-1.4321	42	0.1595	-0.0806	0.0137
sex	-0.0727	0.0426	-1.7066	42	0.0953	-0.1587	0.0133
phase:group	0.0738	0.0405	1.8211	42	0.0757	-0.0080	0.1556
phase:sex	0.0110	0.0469	0.2351	42	0.8152	-0.0836	0.1057
group:sex	0.0434	0.0505	0.8590	42	0.3952	-0.0585	0.1453
phase:group:sex	-0.0437	0.0623	-0.7008	42	0.4873	-0.1695	0.0821
γ =P(Stay|better)							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.6346	0.0143	44.351	42	<.0001	0.6058	0.6635
phase	-0.2521	0.0472	-5.3433	42	<.0001	-0.3473	-0.1569
group	-0.0766	0.0397	-1.9307	42	0.0603	-0.1566	0.0035
sex	-0.1078	0.0637	-1.6919	42	0.0981	-0.2363	0.0208
phase:group	0.0172	0.0791	0.2172	42	0.8291	-0.1424	0.1768
phase:sex	0.1202	0.0834	1.4414	42	0.1569	-0.0481	0.2886
group:sex	0.0493	0.0784	0.6281	42	0.5333	-0.1090	0.2075
phase:group:sex	0.0086	0.1088	0.0787	42	0.9376	-0.2109	0.2281
γ =P(Stay|worse)							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.3761	0.0219	17.134	42	<.0001	0.3318	0.4204
phase	0.2285	0.0434	5.2692	42	<.0001	0.1410	0.3160
group	0.0295	0.0333	0.8868	42	0.3802	-0.0377	0.0968
sex	-0.0157	0.0268	-0.5869	42	0.5604	-0.0697	0.0383
phase:group	0.0267	0.0637	0.4200	42	0.6766	-0.1018	0.1553
phase:sex	-0.0718	0.0497	-1.4435	42	0.1563	-0.1721	0.0286
group:sex	0.0501	0.0437	1.1464	42	0.2581	-0.0381	0.1383
phase:group:sex	-0.0548	0.0754	-0.7274	42	0.4710	-0.2069	0.0973

Table 13. *Perseveration index and repetition index (RI) during early discrimination and reversal learning*

Early Discrimination and Reversal Learning γ~[1+phase*group*sex+(1|rats)]							
γ =Perseveration Index							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.2855	0.0398	7.1683	42	<.0001	0.2052	0.3659
phase	0.3710	0.0836	4.4373	42	**0.0001**	0.2023	0.5398
group	0.0379	0.0461	0.8232	42	0.4151	-0.0551	0.1309
sex	-0.0078	0.0456	-0.1707	42	0.8653	-0.0998	0.0842
phase:group	0.0412	0.1101	0.3745	42	0.7099	-0.1810	0.2635
phase:sex	-0.1251	0.0936	-1.3369	42	0.1885	-0.3139	0.0637
group:sex	0.0424	0.0620	0.6838	42	0.4978	-0.0828	0.1676
phase:group:sex	-0.1286	0.1296	-0.9924	42	0.3267	-0.3902	0.1330
γ =RI							
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.0052	0.0049	1.0620	42	0.2943	-0.0047	0.0152
phase	-0.0109	0.0130	-0.8370	42	0.4074	-0.0371	0.0153
group	-0.0243	0.0106	-2.2810	42	**0.0277**	-0.0457	-0.0028
sex	-0.0623	0.0245	-2.5417	42	**0.0148**	-0.1117	-0.0128
phase:group	0.0247	0.0185	1.3372	42	0.1883	-0.0126	0.0621
phase:sex	0.0245	0.0306	0.8026	42	0.4267	-0.0372	0.0862
group:sex	0.0513	0.0290	1.7718	42	0.0837	-0.0071	0.1098
phase:group:sex	-0.0275	0.0365	-0.7531	42	0.4556	-0.1012	0.0462

Table 14. *Repetition index on the better option (RI_B) and repetition index on the worse option (RI_W) during early discrimination and reversal learning*

Early Discrimination and Reversal Learning γ~[1+phase*group*sex+(1|rats)]							
			$\gamma = RI_B$				
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.0024	0.0023	1.0606	42	0.2949	-0.0022	0.0070
phase	-0.0054	0.0064	-0.8437	42	0.4036	-0.0182	0.0075
group	-0.0121	0.0052	-2.3252	42	**0.0250**	-0.0225	-0.0016
sex	-0.0308	0.0122	-2.5192	42	**0.0157**	-0.0555	-0.0061
phase:group	0.0122	0.0092	1.3214	42	0.1935	-0.0064	0.0308
phase:sex	0.0122	0.0152	0.8033	42	0.4263	-0.0185	0.0429
group:sex	0.0259	0.0145	1.7912	42	0.0805	-0.0033	0.0551
phase:group:sex	-0.0139	0.0182	-0.7607	42	0.4511	-0.0506	0.0229
			$\gamma = RI_W$				
Coefficients	**β**	SE	tStat	DF	p	CI_L	CI_U
intercept	0.0028	0.0027	1.0606	42	0.2949	-0.0025	0.0082
phase	-0.0055	0.0066	-0.8296	42	0.4115	-0.0189	0.0079
group	-0.0122	0.0055	-2.2351	42	**0.0308**	-0.0232	-0.0012
sex	-0.0314	0.0123	-2.5632	42	**0.0140**	-0.0562	-0.0067
phase:group	0.0126	0.0093	1.3515	42	0.1838	-0.0062	0.0313
phase:sex	0.0123	0.0154	0.8016	42	0.4273	-0.0187	0.0433
group:sex	0.0254	0.0145	1.7514	42	0.0872	-0.0039	0.0547
phase:group:sex	-0.0136	0.0183	-0.7451	42	0.4604	-0.0506	0.0233

APPENDIX C: Chapter 4 Computational Modeling and Analyses

RESULTS

Accuracy in reversal learning across session and trials

To gain more insight into this sex difference and its potential underlying mechanism, we next compared the estimated model parameters from the reinforcement learning models (*RL*, RL_{decay}). Comparing the goodness of fit between the two models, we found that the second model with the decay parameter (RL_{decay}) better accounted for the animals' choice behavior as indicated by significantly lower AIC (paired sample t-test; t(1554) = 6.792, $p = 1.56e^{-11}$). Overall mean BIC value was also lower for the second model, although the values did not significantly differ from the first model (t(1554) = 0.93806, $p = 0.348$). Therefore, we focused on the estimated parameters from the $RL1_{decay}$ model only.

Comparison of the estimated parameters across groups revealed that female and male rats differed mainly in the decay parameter γ_d, which governs the amount of passive decay or forgetting in the value estimate of the unchosen option (**Figure C-1**). During the first deterministic reversal (R1, 100/0), eGFP females showed overall significantly lower values of γ_d than eGFP males (mean difference in γ_d = -0.098; Wilcoxon rank sum test, p=0.00537), suggesting a different mechanism of adjustment to the reversal between females and males. While there was no clear evidence for such sex difference in γ_d for the vlOFC or BLA hM4Di groups during the first deterministic reversal (R1), the first probabilistic reversal (R3, 90/10) instead revealed a sex-specific effect between CNO and VEH groups: vlOFC hM4Di females who were administered CNO to inhibit vlOFC had significantly higher values of γ_d compared to those who received VEH (mean difference in γ_d = 0.150; $p = 0.00757$). In contrast, vlOFC hM4Di males did not show a significant difference in γ_d between CNO and VEH groups (mean difference in γ_d = -0.0753; $p = 0.228$). BLA

hM4Di groups showed a similar trend in γ_d, with the females exhibiting a larger difference between CNO and VEH groups (mean difference in γ_d = 0.137; p = 0.0673) than males (mean difference in γ_d = 0.052; p = 0.961). These results based on RL model fitting suggest that the attenuated probabilistic learning for the female CNO groups is mediated by larger γ_d or decreased memory for the unchosen options after vlOFC and BLA inhibition. Importantly, this significant difference emerged due to hM4Di VEH females exhibiting *enhanced* memory in R3. This is because in this condition, rats had previously received CNO in R2 and thus did not encode the reversal very well, making it easier to return to the R1 contingency (which was the same as in R3). Collectively, this shows that vlOFC, and perhaps secondarily BLA, is necessary for the encoding and retrieval of action-based values.

METHODS

Reinforcement Learning Models

To capture the differences in learning and choice behavior during the action-based task, we utilized two conventional reinforcement learning (RL) models. Specifically, the subjective estimate of reward (V) for each choice option was updated on a trial-by-trial basis using reward prediction error (RPE), the discrepancy between actual and expected reward value. In the first model, which we refer to as *RL*, the value estimate of the chosen option (V_C) for a trial t was updated using the following equations:

$$V_C(t+1) = V_C(t) + \alpha\big(R(t) - V_C(t)\big), \qquad (1)$$

where $R(t)$ indicates the presence (1) or absence (0) of a reward for the given trial, and α is the learning rate dictating the amount of update in the value estimate by RPE. In this model, the value of unchosen option was not updated.

The second model, referred to as RL_{decay}, used the same learning rule as Equation (1) for updating the value of the chosen option, and additionally updated the value of the unchosen option (V_U) as follows:

$$V_U(t+1) = V_U(t) - \gamma_d(V_U(t)), \qquad (2)$$

where γ_d is a decay rate controlling the amount of passive decay in value of the unchosen option. In both models above, the probability of choosing a particular option was computed using the following decision rule:

$$P_i(t) = (1 + e^{-\beta\left(V_i(t)-V_j(t)\right)})^{-1}, \qquad (3)$$

where *i* and *j* corresponds to two alternative options (i.e., left and right for action-based task), and β is the inverse temperature or sensitivity governing the extent to which higher-valued options are consistently selected.

We used the standard maximum likelihood estimation method to fit choice data and estimate the parameters for each session of the experiment. The values of the learning rate α and decay rate γ_d were bounded between 0 and 1, and β was bounded between 1 and 100. Initial parameter values were selected from this range, and fitting was performed using the MATLAB function *fmincon*. For each set of parameters fitted to each session, we repeated 30 different initial conditions selected from evenly-spaced search space to avoid local minima. The best fit was selected from the iteration with the minimum negative log-likelihood (*LL*). For the first model (*RL*), we treated the uninitiated or uncommitted trials with no choice data as if they had not occurred. In contrast, for the second model (RL_{decay}), both choice options were considered unchosen for those trials and both of the value estimates decayed passively according to Equation (2).

To quantify goodness of fit, we computed both Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for each session as follows:

$$AIC = -2 * LL + 2 * k, \quad (4)$$

$$BIC = -2 * LL + \ln(n) * k, \quad (5)$$

where k is the number of parameters in the model (two for RL and three for RL_{decay}), and n is the number of choice trials in the session.

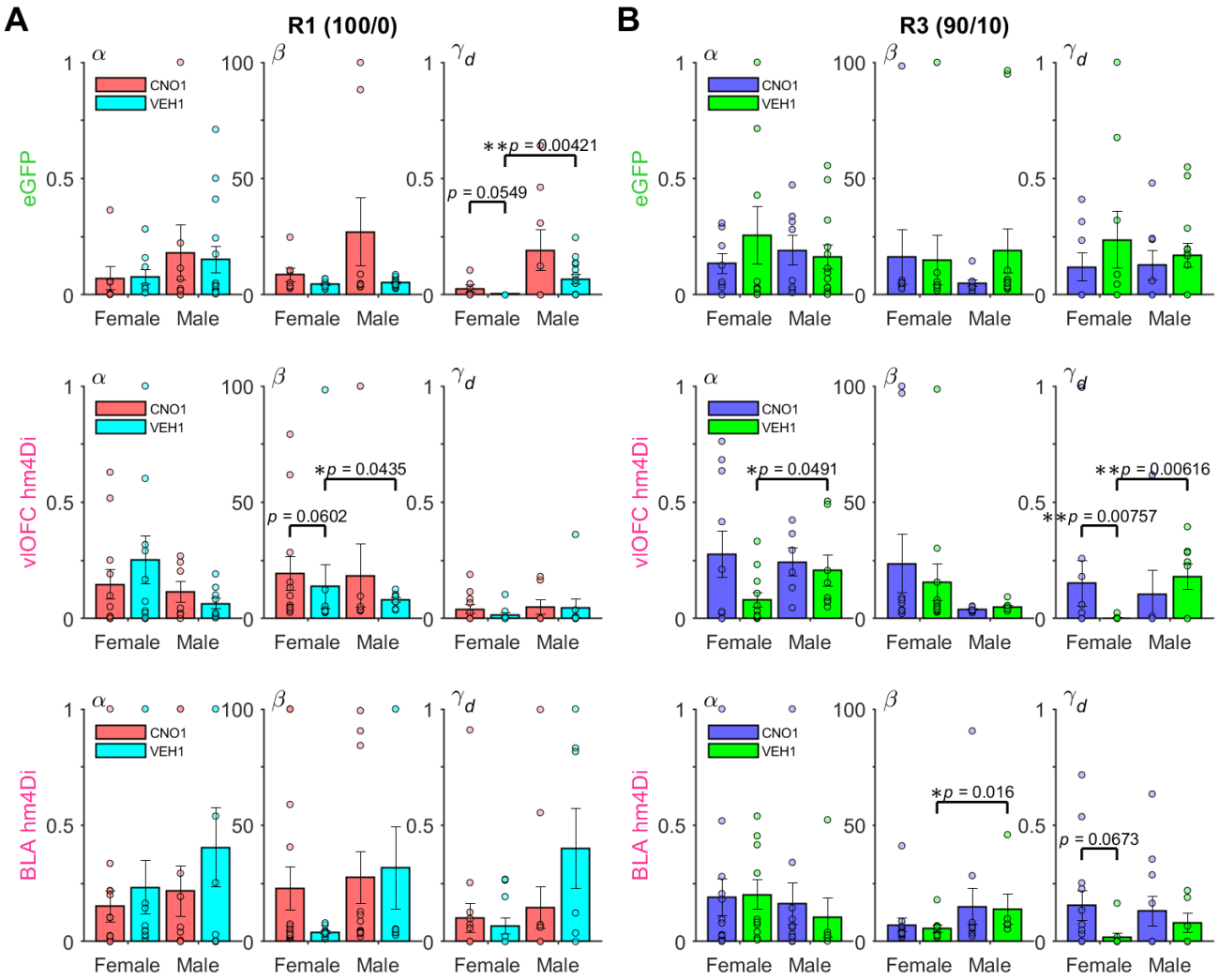

Figure C-1. Reinforcement learning model fit to choice behavior indicates differential effect on the decay parameter between male and female rats during deterministic and probabilistic action-based reversals. Plotted are estimated model parameters for $RL1_{decay}$ fitted to each session during deterministic (100/0) **(A)** and probabilistic (90/10) **(B)** reversals. α = learning rate, β = inverse temperature, γ_d = decay rate for unchosen option. Drug order was counterbalanced such that on R2 and R4 (not shown) animals received VEH if they were administered CNO first on R1 and R3, and vice versa. Female eGFP group exhibited overall

References

Aarde, S. M., Genner, R. M., Hrncir, H., Arnold, A. P., & Jentsch, J. D. (2020). Sex chromosome complement affects multiple aspects of reversal-learning task performance in mice. *Genes, Brain and Behavior.* e12685. https://doi.org/10.1111/gbb.12685

Ajayi, A. F., & Akhigbe, R. E. (2020). Staging of the estrous cycle and induction of estrus in experimental rodents: an update. *Fertility research and practice*, *6*, 5. https://doi.org/10.1186/s40738-020-00074-3

Almeida OF, Shoaib M, Deicke J, Fischer D, Darwish MH, Patchev VK (1998) Gender differences in ethanol preference and ingestion in rats. The role of the gonadal steroid environment. *J Clin Invest* 101:2677–2685.

Aguirre, C. G., Stolyarova, A., Das, K., Kolli, S., Marty, V., Ray, L., Spigelman, I., & Izquierdo, A. (2020). Sex-dependent effects of chronic intermittent voluntary alcohol consumption on attentional, not motivational, measures during probabilistic learning and reversal. *PLOS ONE*, *15*(6), e0234729. https://doi.org/10.1371/journal.pone.0234729

Alsiö, J., Nilsson, S. R. O., Gastambide, F., Wang, R. A. H., Dam, S. A., Mar, A. C., Tricklebank, M., & Robbins, T. W. (2015). The role of 5-HT2C receptors in touchscreen visual reversal learning in the rat: A cross-site study. *Psychopharmacology*, *232*(21), 4017–4031. https://doi.org/10.1007/s00213-015-3963-5

Alvarez, P., & Eichenbaum, H. (2002). Representations of odors in the rat orbitofrontal cortex change during and after learning. *Behavioral Neuroscience*, *116*(3), 421–433. https://doi.org/10.1037/0735-7044.116.3.421

Amitai, N., & Markou, A. (2010). Disruption of performance in the five-choice serial reaction time task induced by administration of N-methyl-D-aspartate receptor antagonists: relevance to

cognitive dysfunction in schizophrenia. *Biol Psychiatry*, *68*(1), 5-16. https://doi.org/10.1016/j.biopsych.2010.03.004

Amodeo, L. R., McMurray, M. S., & Roitman, J. D. (2017). Orbitofrontal cortex reflects changes in response–outcome contingencies during probabilistic reversal learning. *Cognitive Flexibility: Development, Disease, and Treatment*, *345*, 27–37. https://doi.org/10.1016/j.neuroscience.2016.03.034

Amsel, A. (Ed.). (1967). *Partial reinforcement effects in vigor and persistence: Advances in frustration theory derived from a variety of within-subjects experiments.* (Vol. 1). Academic Press.

Asinof, S. K., & Paine, T. A. (2014). The 5-choice serial reaction time task: a task of attention and impulse control for rodents. *J Vis Exp*(90), e51574. https://doi.org/10.3791/51574

Asrican, B., & Song, J. (2021). Extracting meaningful circuit-based calcium dynamics in astrocytes and neurons from adult mouse brain slices using single-photon GCaMP imaging. *STAR Protoc*, *2*(1), 100306. https://doi.org/10.1016/j.xpro.2021.100306

Averbeck, B.B. (2021). Effects of Amygdala Lesions on Object-Based Versus Action-Based Learning in Macaques. *Cereb Cortex* **31**(1): 529-546.

Averbeck, B. B., & Costa, V. D. (2017). Motivational neural circuits underlying reinforcement learning. *Nature Neuroscience*, *20*(4), 505–512. https://doi.org/10.1038/nn.4506

Badanich, K. A., Fakih, M. E., Gurina, T. S., Roy, E. K., Hoffman, J. L., Uruena-Agnes, A. R., & Kirstein, C. L. (2016). Reversal learning and experimenter-administered chronic intermittent ethanol exposure in male rats. *Psychopharmacology*, *233*(19), 3615–3626. https://doi.org/10.1007/s00213-016-4395-6

Badanich, K. A., Mulholland, P. J., Beckley, J. T., Trantham-Davidson, H., & Woodward, J. J. (2013). Ethanol Reduces Neuronal Excitability of Lateral Orbitofrontal Cortex Neurons Via a Glycine Receptor Dependent Mechanism. *Neuropsychopharmacology*, *38*, 1176.

Badanich, K. A., Becker, H. C., & Woodward, J. J. (2011). Effects of chronic intermittent ethanol exposure on orbitofrontal and medial prefrontal cortex-dependent behaviors in mice. *Behavioral Neuroscience*, *125*(6), 879–891. PMC. https://doi.org/10.1037/a0025922

Balleine, B. W., Leung, B. K., & Ostlund, S. B. (2011). The orbitofrontal cortex, predicted value, and choice. *Annals of the New York Academy of Sciences*, *1239*(1), 43–50. https://doi.org/10.1111/j.1749-6632.2011.06270.x

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: An integrated view of amygdala function. *Trends in Neurosciences*, *29*(5), 272–279. https://doi.org/10.1016/j.tins.2006.03.002

Balleine, B. W., Killcross, A. S., & Dickinson, A. (2003). The Effect of Lesions of the Basolateral Amygdala on Instrumental Conditioning. *The Journal of Neuroscience*, *23*(2), 666. https://doi.org/10.1523/JNEUROSCI.23-02-00666.2003

Baltz, E. T., Yalcinbas, E. A., Renteria, R., & Gremel, C. M. (2018). Orbital frontal cortex updates state-induced value change for decision-making. *eLife*, *7*, e35988. https://doi.org/10.7554/eLife.35988

Bari, A., Dalley, J. W., & Robbins, T. W. (2008). The application of the 5-choice serial reaction time task for the assessment of visual attentional processes and impulse control in rats. *Nat Protoc*, *3*(5), 759-767. https://doi.org/10.1038/nprot.2008.41

Barker, J. M., & Taylor, J. R. (2014). Habitual alcohol seeking: Modeling the transition from casual drinking to addiction. *Neuroscience & Biobehavioral Reviews*, *47*, 281-294. https://doi.org/https://doi.org/10.1016/j.neubiorev.2014.08.012

Barker JM, Torregrossa MM, Arnold AP, Taylor JR. (2010). Dissociation of genetic and hormonal influences on sex differences in alcoholism-related behaviors. *J Neurosci.* 30(27):9140–4.

Barreiros, I. V., Ishii, H., Walton, M. E., & Panayi, M. C. (2021). Defining an orbitofrontal compass: Functional and anatomical heterogeneity across anterior-posterior and medial-lateral axes. *Behav Neurosci*, *135*(2), 165-173. https://doi.org/10.1037/bne0000442

Barreiros, I. V., Panayi, M. C., & Walton, M. E. (2021). Organization of Afferents along the Anterior-posterior and Medial-lateral Axes of the Rat Orbitofrontal Cortex. *Neuroscience*, *460*, 53-68. https://doi.org/10.1016/j.neuroscience.2021.02.017

Barth, C., Villringer, A., & Sacher, J. (2015). Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods. *Frontiers in neuroscience*, *9*, 37. https://doi.org/10.3389/fnins.2015.00037

Bathellier, B., Tee, S. P., Hrovat, C., & Rumpel, S. (2013). A multiplicative reinforcement learning model capturing learning dynamics and interindividual variability in mice. *Proc Natl Acad Sci U S A*, *110*(49), 19950-19955. https://doi.org/10.1073/pnas.1312125110

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

Bayless DW, Darling JS, Stout WJ, Daniel JM. (2012). Sex differences in attentional processes in adult rats as measured by performance on the 5-choice serial reaction time task. *Behavioural Brain Research*. 235(1):48–54.

Becker, J. B., Prendergast, B. J., & Liang, J. W. (2016). Female rats are not more variable than male rats: a meta-analysis of neuroscience studies. *Biology of sex differences*, *7*, 34. https://doi.org/10.1186/s13293-016-0087-5

Becker HC, Ron D. (2014).. Animal models of excessive alcohol consumption: Recent advances and future challenges. *Alcohol.* 48(3):205–8.

Becker, J. B., Arnold, A. P., Berkley, K. J., Blaustein, J. D., Eckel, L. A., Hampson, E., Herman, J. P., Marts, S., Sadee, W., Steiner, M., Taylor, J., & Young, E. (2005). Strategies and methods for research on sex differences in brain and behavior. *Endocrinology*, *146*(4), 1650–1673. https://doi.org/10.1210/en.2004-1142

Beery A. K. (2018). Inclusion of females does not increase variability in rodent research studies. *Current opinion in behavioral sciences*, *23*, 143–149. https://doi.org/10.1016/j.cobeha.2018.06.016

Beery, A. K. and I. Zucker (2011). Sex bias in neuroscience and biomedical research. *Neuroscience and biobehavioral reviews* **35**(3): 565-572.

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nat Neurosci*, *10*(9), 1214-1221. https://doi.org/10.1038/nn1954

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Bissonette, G. B., Lande, M. D., Martins, G. J., & Powell, E. M. (2012). Versatility of the mouse reversal/set-shifting test: Effects of topiramate and sex. *International Behavioral Neuroscience Society* (IBNS), 107(5), 781–786. https://doi.org/10.1016/j.physbeh.2012.05.018

Bolla, K. I., D. A. Eldreth, J. A. Matochik and J. L. Cadet (2004). Sex-related Differences in a

Gambling Task and Its Neurological Correlates. *Cerebral Cortex* **14**(11): 1226-1232.

Boulougouris, V., & Robbins, T. W. (2009). Pre-surgical training ameliorates orbitofrontal-mediated impairments in spatial reversal learning. *Behavioural Brain Research*, *197*(2), 469–475. https://doi.org/10.1016/j.bbr.2008.10.005

Boulougouris, V., Dalley, J. W., & Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research*, *179*(2), 219–228. https://doi.org/10.1016/j.bbr.2007.02.005

Boutros N, Der-Avakian A, Markou A, Semenova S. (2017). Effects of early life stress and adolescent ethanol exposure on adult cognitive performance in the 5-choice serial reaction time task in Wistar male rats. *Psychopharmacology*. 234(9):1549–56.

Branch, C. L., Sonnenberg, B. R., Pitera, A. M., Benedict, L. M., Kozlovsky, D. Y., Bridge, E. S., & Pravosudov, V. V. (2020). Testing the greater male variability phenomenon: Male mountain chickadees exhibit larger variation in reversal learning performance compared with females. *Proceedings of the Royal Society B: Biological Sciences*, 287(1931), 20200895. https://doi.org/10.1098/rspb.2020.0895

Brigman, J. L., Feyder, M., Saksida, L. M., Bussey, T. J., Mishina, M., & Holmes, A. (2008). Impaired discrimination learning in mice lacking the NMDA receptor NR2A subunit. *Learn Mem*, *15*(2), 50-54. https://doi.org/10.1101/lm.777308

Brigman, J. L., Graybeal, C., & Holmes, A. (2010). Predictably irrational: assaying cognitive inflexibility in mouse models of schizophrenia. *Frontiers in neuroscience*, *4*. https://doi.org/10.3389/neuro.01.013.2010

Brigman, J. L., Mathur, P., Harvey-White, J., Izquierdo, A., Saksida, L. M., Bussey, T. J., Fox, S., Deneris, E., Murphy, D. L., & Holmes, A. (2010). Pharmacological or genetic inactivation

of the serotonin transporter improves reversal learning in mice. *Cereb Cortex*, *20*(8), 1955-1963. https://doi.org/10.1093/cercor/bhp266

Bruinsma B, Terra H, de Kloet SF, Luchicchi A, Timmerman AJ, Remmelink E. (2019). An automated home-cage-based 5-choice serial reaction time task for rapid assessment of attention and impulsivity in rats. *Psychopharmacology* (Berl). 236(7):2015–26.

Brushfield, A. M., Luu, T. T., Callahan, B. D., & Gilbert, P. E. (2008). A comparison of discrimination and reversal learning for olfactory and visual stimuli in aged rats. *Behavioral Neuroscience*, *122*(1), 54–62. PubMed. https://doi.org/10.1037/0735-7044.122.1.54

Bryce, C. A., & Floresco, S. B. (2021). Central CRF and acute stress differentially modulate probabilistic reversal learning in male and female rats. *Behavioural brain research*, *397*, 112929. https://doi.org/10.1016/j.bbr.2020.112929

Bryce CA, Howland JG. (2015). Stress facilitates late reversal learning using a touchscreen-based visual discrimination procedure in male Long Evans rats. *Behavioural Brain Research.* 278:21–8.

Brys I, Pupe S, Bizarro L. (2014). Attention, locomotor activity and developmental milestones in rats prenatally exposed to ethanol. *International Journal of Developmental Neuroscience*. 38:161–8.

Bushnell, P. J., & Strupp, B. J. (2009). Assessing Attention in Rodents. In nd & J. J. Buccafusco (Eds.), *Methods of Behavior Analysis in Neuroscience*. https://www.ncbi.nlm.nih.gov/pubmed/21204340

Cailhol S, Mormède P. (2002) Conditioned taste aversion and alcohol drinking: strain and gender differences. *J Stud Alcohol.* 63(1):91–9.

Carlsen, J. (1988). Immunocytochemical localization of glutamate decarboxylase in the rat basolateral amygdaloid nucleus, with special reference to GABAergic innervation of amygdalostriatal projection neurons. *Journal of Comparative Neurology*, *273*(4), 513–526. https://doi.org/10.1002/cne.902730407

Carnicella S, Ron D, Barak S. (2014). Intermittent ethanol access schedule in rats as a preclinical model of alcohol abuse. *Alcohol*. 48(3):243–52.

Carnicella S, Amamoto R, Ron D. (2009). Excessive alcohol consumption is blocked by glial cell line–derived neurotrophic factor. *Alcohol.* 43(1):35–43

Carroll, H. A., Lustyk, M. K., & Larimer, M. E. (2015). The relationship between alcohol consumption and menstrual cycle: a review of the literature. *Archives of women's mental health*, *18*(6), 773–781. https://doi.org/10.1007/s00737-015-0568-2

Cassell, M. D., & Wright, D. J. (1986). Topography of projections from the medial prefrontal cortex to the amygdala in the rat. *Brain Research Bulletin*, *17*(3), 321–333. https://doi.org/10.1016/0361-9230(86)90237-6

Charlton, A. J., May, C., Luikinga, S. J., Burrows, E. L., Hyun Kim, J., Lawrence, A. J., & Perry, C. J. (2019). Chronic voluntary alcohol consumption causes persistent cognitive deficits and cortical cell loss in a rodent model. *Scientific Reports*, *9*(1), 18651. https://doi.org/10.1038/s41598-019-55095-w

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. *Curr Biol*, *31*(1), 39-50 e34. 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*. https://doi.org/10.7554/eLife.69748

Chudasama, Y., Passetti, F., Rhodes, S. E. V., Lopian, D., Desai, A., & Robbins, T. W. (2003). Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: Differential effects on selectivity, impulsivity and compulsivity. *The Rodent Prefrontal Cortex*, *146*(1), 105–119. https://doi.org/10.1016/j.bbr.2003.09.020

Chudasama, Y., & Robbins, T. W. (2003). Dissociable Contributions of the Orbitofrontal and Infralimbic Cortex to Pavlovian Autoshaping and Discrimination Reversal Learning: Further Evidence for the Functional Heterogeneity of the Rodent Frontal Cortex. *The Journal of Neuroscience*, *23*(25), 8771. https://doi.org/10.1523/JNEUROSCI.23-25-08771.2003

Churchwell, J. C., Morris, A. M., Heurtelou, N. M., & Kesner, R. P. (2009). Interactions between the prefrontal cortex and amygdala during delay discounting and reversal. *Behavioral Neuroscience*, *123*(6), 1185–1196. PubMed. https://doi.org/10.1037/a0017734

Coleman, L. G., Jr, Liu, W., Oguz, I., Styner, M., & Crews, F. T. (2014). Adolescent binge ethanol treatment alters adult brain regional volumes, cortical extracellular matrix protein and behavioral flexibility. *Pharmacology, Biochemistry, and Behavior*, *116*, 142–151. PubMed. https://doi.org/10.1016/j.pbb.2013.11.021

Coleman, L. G., Jr., He, J., Lee, J., Styner, M., & Crews, F. T. (2011). Adolescent binge drinking alters adult brain neurotransmitter gene expression, behavior, brain regional volumes, and neurochemistry in mice. *Alcohol Clin Exp Res*, *35*(4), 671-688. https://doi.org/10.1111/j.1530-0277.2010.01385.x

Corbit, L. H., & Balleine, B. W. (2005). Double Dissociation of Basolateral and Central Amygdala Lesions on the General and Outcome-Specific Forms of Pavlovian-

Instrumental Transfer. *The Journal of Neuroscience*, *25*(4), 962. https://doi.org/10.1523/JNEUROSCI.4507-04.2005

Costa, K. M., Scholz, R., Lloyd, K., Moreno-Castilla, P., Gardner, M. P. H., Dayan, P., & Schoenbaum, G. (2023). The role of the lateral orbitofrontal cortex in creating cognitive maps. *Nat Neurosci*, *26*(1), 107-115. https://doi.org/10.1038/s41593-022-01216-0

Costa, V. D., Dal Monte, O., Lucas, D. R., Murray, E. A., & Averbeck, B. B. (2016). Amygdala and Ventral Striatum Make Distinct Contributions to Reinforcement Learning. *Neuron*, *92*(2), 505–517. https://doi.org/10.1016/j.neuron.2016.09.025

Costa, V. D., Tran, V. L., Turchi, J., & Averbeck, B. B. (2015). Reversal learning and dopamine: a bayesian perspective. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *35*(6), 2407-2416. https://doi.org/10.1523/jneurosci.1989-14.2015

Cox, J., A. R. Minerva, W. T. Fleming, C. A. Zimmerman, C. Hayes, S. Zorowitz, A. Bandi, S. Ornelas, B. McMannon, N. F. Parker and I. B. Witten (2023). A neural substrate of sex-dependent modulation of motivation. *Nature Neuroscience* **26**(2): 274-284.

Crabbe, J. C., Harris, R. A., & Koob, G. F. (2011). Preclinical studies of alcohol binge drinking. *Ann N Y Acad Sci*, *1216*, 24-40. https://doi.org/10.1111/j.1749-6632.2010.05895.x

Crews, F. T., & Boettiger, C. A. (2009). Impulsivity, frontal lobes and risk for addiction. *Pharmacol Biochem Behav*, *93*(3), 237-247. https://doi.org/10.1016/j.pbb.2009.04.018

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends in Neurosciences*, *38*(9), 571–578. PubMed. https://doi.org/10.1016/j.tins.2015.07.003

Dalton, G. L., Wang, N. Y., Phillips, A. G., & Floresco, S. B. (2016). Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic

Reversal Learning. *The Journal of Neuroscience*, *36*(6), 1996. https://doi.org/10.1523/JNEUROSCI.3366-15.2016

Dalton, G. L., Phillips, A. G., & Floresco, S. B. (2014). Preferential involvement by nucleus accumbens shell in mediating probabilistic learning and reversal shifts. *The Journal of Neuroscience*, *34*(13), 4618–4626. https://doi.org/10.1523/JNEUROSCI.5058-13.2014

Daviu, N., Bruchas, M. R., Moghaddam, B., Sandi, C., & Beyeler, A. (2019). Neurobiological links between stress and anxiety. *Neurobiology of Stress*, *11*, 100191. https://doi.org/10.1016/j.ynstr.2019.100191

Daw, N. D. (2011). *Trial-by-trial data analysis using computational models* (Vol. 23). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199600434.003.0001

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 473–492. https://doi.org/10.3758/s13415-014-0277-8

Dayan, P., & Niv, Y. (2008). Reinforcement learning: The Good, The Bad and The Ugly. *Cognitive Neuroscience*, *18*(2), 185–196. https://doi.org/10.1016/j.conb.2008.08.003

de la Torre, M. L., Escarabajal, M. D., & Agüero, Á. (2015). Sex differences in adult Wistar rats in the voluntary consumption of ethanol after pre-exposure to ethanol-induced flavor avoidance learning. *Pharmacol Biochem Behav*, *137*, 7-15. https://doi.org/10.1016/j.pbb.2015.07.011

De Steno, D. A., & Schmauss, C. (2009). A role for dopamine D2 receptors in reversal learning. *Neuroscience*, *162*(1), 118–127. https://doi.org/10.1016/j.neuroscience.2009.04.052

Del Río, J. P., Alliende, M. I., Molina, N., Serrano, F. G., Molina, S., & Vigil, P. (2018).

Steroid Hormones and Their Action in Women's Brains: The Importance of Hormonal Balance. *Frontiers in public health*, *6*, 141. https://doi.org/10.3389/fpubh.2018.00141

DePoy L, Daut R, Brigman JL, MacPherson K, Crowley N, Gunduz-Cinar O. (2013).. Chronic alcohol produces neuroadaptations to prime dorsal striatal learning. *Proc Natl Acad Sci*. 110(36):14783.

Derman, R. C., Bass, C. E., & Ferrario, C. R. (2020). Effects of hM4Di activation in CamKII basolateral amygdala neurons and CNO treatment on sensory-specific vs. General PIT: refining PIT circuits and considerations for using CNO. *Psychopharmacology*, *237*(5), 1249–1266. https://doi.org/10.1007/s00213-020-05453-8

Dreher, J.-C., P. J. Schmidt, P. Kohn, D. Furman, D. Rubinow and K. F. Berman (2007). Menstrual cycle phase modulates reward-related neural function in women. *Proceedings of the National Academy of Sciences of the United States of America* **104**(7): 2465-2470.

Eichenbaum, H., Fagan, A., & Cohen, N. (1986). Normal olfactory discrimination learning set and facilitation of reversal learning after medial-temporal damage in rats: Implications for an account of preserved learning abilities in amnesia. *The Journal of Neuroscience*, *6*(7), 1876. https://doi.org/10.1523/JNEUROSCI.06-07-01876.1986

Erol, A., Ho, A. M., Winham, S. J., & Karpyak, V. M. (2019). Sex hormones in alcohol consumption: a systematic review of evidence. *Addict Biol*, *24*(2), 157-169. https://doi.org/10.1111/adb.12589

Esber, G. R., M. R. Roesch, S. Bali, J. Trageser, G. B. Bissonette, A. C. Puche, P. C. Holland and G. Schoenbaum (2012). Attention-related Pearce-Kaye-Hall signals in basolateral amygdala require the midbrain dopaminergic system. *Biological psychiatry* **72**(12): 1012-1019.

Evans KL, Hampson E. (2015). Sex differences on prefrontally-dependent cognitive tasks. *Brain and Cognition*. 93:42–53.

Fanselow, M., & Gale, G. (2003). The Amygdala, Fear, and Memory. *Annals of the New York Academy of Sciences*, *985*, 125–134. https://doi.org/10.1111/j.1749-6632.2003.tb07077.x

Fanselow, M. S., & LeDoux, J. E. (1999). Why We Think Plasticity Underlying Pavlovian Fear Conditioning Occurs in the Basolateral Amygdala. *Neuron*, *23*(2), 229–232. https://doi.org/10.1016/S0896-6273(00)80775-8

Farashahi, S., Donahue, C. H., Khorsand, P., Seo, H., Lee, D., & Soltani, A. (2017). Metaplasticity as a Neural Substrate for Adaptive Learning and Choice under Uncertainty. *Neuron*, *94*(2), 401-414.e6. https://doi.org/10.1016/j.neuron.2017.03.044

Farrell, M. R., Gruene, T. M., & Shansky, R. M. (2015). The influence of stress and gonadal hormones on neuronal structure and function. *Hormones and behavior*, *76*, 118–124. https://doi.org/10.1016/j.yhbeh.2015.03.003

Fernandez, G. M., Lew, B. J., Vedder, L. C., & Savage, L. M. (2017). Chronic intermittent ethanol exposure leads to alterations in brain-derived neurotrophic factor within the frontal cortex and impaired behavioral flexibility in both adolescent and adult rats. *Neuroscience*, *348*, 324–334. https://doi.org/10.1016/j.neuroscience.2017.02.045

Fernandez, G. M., Stewart, W. N., & Savage, L. M. (2016). Chronic Drinking During Adolescence Predisposes the Adult Rat for Continued Heavy Drinking: Neurotrophin and Behavioral Adaptation after Long-Term, Continuous Ethanol Exposure. *PLoS ONE*, *11*(3), e0149987. PMC. https://doi.org/10.1371/journal.pone.0149987

Fidler, T. L., Powers, M. S., Ramirez, J. J., Crane, A., Mulgrew, J., Smitasin, P., & Cunningham, C. L. (2012). Dependence induced increases in intragastric alcohol consumption in mice.

Addiction Biology, *17*(1), 13-32. https://doi.org/https://doi.org/10.1111/j.1369-1600.2011.00363.x

Fineberg, N. A., Potenza, M. N., Chamberlain, S. R., Berlin, H. A., Menzies, L., Bechara, A., Sahakian, B. J., Robbins, T. W., Bullmore, E. T., & Hollander, E. (2010). Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review. *Neuropsychopharmacology*, *35*(3), 591-604. https://doi.org/10.1038/npp.2009.185

Finger, E. C., Marsh, A. A., Mitchell, D. G., Reid, M. E., Sims, C., Budhani, S., Kosson, D. S., Chen, G., Towbin, K. E., Leibenluft, E., Pine, D. S., & Blair, J. R. (2008). Abnormal ventromedial prefrontal cortex function in children with psychopathic traits during reversal learning. *Arch Gen Psychiatry*, *65*(5), 586-594. https://doi.org/10.1001/archpsyc.65.5.586

Finn D. A. (2020). The Endocrine System and Alcohol Drinking in Females. *Alcohol research : current reviews*, *40*(2), 02. https://doi.org/10.35946/arcr.v40.2.02

Fisher, H., Bright, N., Gallo, M., Pajser, A., & Pickens, C. L. (2017). Relationship of low doses of alcohol voluntarily consumed during adolescence and early adulthood with subsequent behavioral flexibility. *Behavioural Pharmacology*, *28*(7). https://journals.lww.com/behaviouralpharm/Fulltext/2017/10000/Relationship_of_low_doses_of_alcohol_voluntarily.4.aspx

Ford MM, Eldridge JC, Samson HH (2004) Determination of an estradiol dose-response relationship in the modulation of ethanol intake. *Alcohol Clin Exp Res* 28:20–28.

Ford MM, Eldridge JC, Samson HH (2002a) Ethanol consumption in the female long-evans rat: a modulatory role of estradiol. *Alcohol* 26:103–113.

Ford MM, Eldridge JC, Samson HH (2002b) Microanalysis of ethanol self-administration:

estrous cycle phase-related changes in consumption patterns. *Alcohol Clin Exp Re*s 26:635–643.

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

Gargiulo, A. T., Hu, J., Ravaglia, I. C., Hawks, A., Li, X., Sweasy, K., & Grafe, L. (2022). Sex differences in cognitive flexibility are driven by the estrous cycle and stress-dependent. *Frontiers in behavioral neuroscience*, *16*, 958301. https://doi.org/10.3389/fnbeh.2022.958301

George O, Sanders C, Freiling J, Grigoryan E, Vu S, Allen CD. (2012) Recruitment of medial prefrontal cortex neurons during alcohol withdrawal predicts cognitive impairment and excessive alcohol drinking. *Proc Natl Acad Sci.* 109(44):18156.

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 : official publication of the American College of Neuropsychopharmacology*, *43*(2), 313–324. https://doi.org/10.1038/npp.2017.161

Gershman, S. J. (2018). The Successor Representation: Its Computational Logic and Neural Substrates. *The Journal of Neuroscience*, *38*(33), 7193. https://doi.org/10.1523/JNEUROSCI.0151-18.2018

Ghasemi, A., Jeddi, S., & Kashfi, K. (2021). The laboratory rat: Age and body weight matter. *EXCLI J*, *20*, 1431-1445. https://doi.org/10.17179/excli2021-4072

Ghashghaei, H. T., & Barbas, H. (2002). Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience*, *115*(4), 1261–1279. https://doi.org/10.1016/S0306-4522(02)00446-3

Givens B. (1997). Effect of ethanol on sustained attention in rats. *Psychopharmacology.* 129(2):135–40.

Goudriaan, A. E., Oosterlaan, J., de Beurs, E., & van den Brink, W. (2006). Neurocognitive functions in pathological gambling: a comparison with alcohol dependence, Tourette syndrome and normal controls. *Addiction*, *101*(4), 534-547. https://doi.org/10.1111/j.1360-0443.2006.01380.x

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nature Communications*, *4*(1), 2264. https://doi.org/10.1038/ncomms3264

Griffin, W. C., 3rd. (2014). Alcohol dependence and free-choice drinking in mice. *Alcohol*, *48*(3), 287-293. https://doi.org/10.1016/j.alcohol.2013.11.006

Grissom NM, Reyes TM. (2019). Let's call the whole thing off: evaluating gender and sex differences in executive function. *Neuropsychopharmacology.* 44(1):86–96.

Groenewegen, H. J. (1988). Organization of the afferent connections of the mediodorsal thalamic nucleus in the rat, related to the mediodorsal-prefrontal topography. *Neuroscience*, *24*(2), 379–431. https://doi.org/10.1016/0306-4522(88)90339-9

Groman, S. M., Keistler, C., Keip, A. J., Hammarlund, E., DiLeone, R. J., Pittenger, C., Lee, D., & Taylor, J. R. (2019). Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes. *Neuron*, *103*(4), 734-746.e3. https://doi.org/10.1016/j.neuron.2019.05.042

Groman, S. M., & Jentsch, J. D. (2013). Identifying the molecular basis of inhibitory control deficits in addictions: neuroimaging in non-human primates. *Curr Opin Neurobiol*, *23*(4), 625-631. https://doi.org/10.1016/j.conb.2013.03.001

Groman, S. M., James, A. S., & Jentsch, J. D. (2009). Poor response inhibition: at the nexus between substance abuse and attention deficit/hyperactivity disorder. *Neurosci Biobehav Rev*, *33*(5), 690-698. https://doi.org/10.1016/j.neubiorev.2008.08.008

Guidi M, Kumar A, Foster TC. (2015). Impaired attention and synaptic senescence of the prefrontal cortex involves redox regulation of NMDA receptors. *J Neurosci.* 35(9):3966–77.

Hamid, A. A. et al. Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* 19, 117–126 (2016).

Hargreaves, G. A., Monds, L., Gunasekaran, N., Dawson, B., & McGregor, I. S. (2009). Intermittent access to beer promotes binge-like drinking in adolescent but not adult Wistar rats. *Alcohol*, *43*(4), 305–314. https://doi.org/10.1016/j.alcohol.2009.02.005

Harris, C., Aguirre, C., Kolli, S., Das, K., Izquierdo, A., & Soltani, A. (2021). Unique features of stimulus-based probabilistic reversal learning. *Behavioral Neuroscience*, *135*(4), 550–570. https://doi.org/10.1037/bne0000474

Hart, E. E., Blair, G. J., O'Dell, T. J., Blair, H. T., & Izquierdo, A. (2020). Chemogenetic Modulation and Single-Photon Calcium Imaging in Anterior Cingulate Cortex Reveal a Mechanism for Effort-Based Decisions. *J Neurosci*, *40*(29), 5628-5643. https://doi.org/10.1523/JNEUROSCI.2548-19.2020

Hervig, M. E., Fiddian, L., Piilgaard, L., Božič, T., Blanco-Pozo, M., Knudsen, C., Olesen, S. F., Alsiö, J., & Robbins, T. W. (2019). Dissociable and Paradoxical Roles of Rat Medial and

Lateral Orbitofrontal Cortex in Visual Serial Reversal Learning. *Cerebral Cortex*, *bhz144*. https://doi.org/10.1093/cercor/bhz144

Hoover, W. B., & Vertes, R. P. (2011). Projections of the medial orbital and ventral orbital cortex in the rat. *Journal of Comparative Neurology*, *519*(18), 3766–3801. https://doi.org/10.1002/cne.22733

Houston RJ, Derrick J, Leonard K, Testa M, Quigley B, Kubiak A. (2014). Effects of Heavy Drinking on Executive Cognitive Functioning in a Community Sample. *Addictive behaviors*. 39(1):345–349.

Hwa, L. S., Chu, A., Levinson, S. A., Kayyali, T. M., DeBold, J. F., & Miczek, K. A. (2011). Persistent escalation of alcohol drinking in C57BL/6J mice with intermittent access to 20% ethanol. *Alcoholism, Clinical and Experimental Research*, *35*(11), 1938–1947. PMC. https://doi.org/10.1111/j.1530-0277.2011.01545.x

Hwang, W. J., Lee, T. Y., Kim, N. S., & Kwon, J. S. (2020). The Role of Estrogen Receptors and Their Signaling across Psychiatric Disorders. *International journal of molecular sciences*, *22*(1), 373. https://doi.org/10.3390/ijms22010373

Iordanova, M. D., J. O.-Y. Yau, M. A. McDannald and L. H. Corbit (2021). Neural substrates of appetitive and aversive prediction error. *Neuroscience and biobehavioral reviews* **123**: 337-351.

IBM SPSS Statistics. Armonk, NY: IBM Corp.; 2017.

Irimia C, Tuong RN, Quach T, Parsons LH. (2014). Impaired Response Inhibition in the Rat 5 Choice Continuous Performance Task during Protracted Abstinence from Chronic Alcohol Consumption. *PLOS ONE.* 9(10):e109948.

Irimia C, Wiskerke J, Natividad LA, Polis IY, Vries TJ, Pattij T. (2015). Increased impulsivity in rats as a result of repeated cycles of alcohol intoxication and abstinence. *Addiction Biology*. 20(2):263–74.

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

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Cognitive Flexibility: Development, Disease, and Treatment*, *345*, 12–26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Izquierdo, A. (2017). Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *37*(44), 10529-10540. https://doi.org/10.1523/jneurosci.1678-17.2017

Izquierdo, A., Darling, C., Manos, N., Pozos, H., Kim, C., Ostrander, S., Cazares, V., Stepp, H., & Rudebeck, P. H. (2013). Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(9), 4105–4109. PubMed. https://doi.org/10.1523/JNEUROSCI.4942-12.2013

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

Izquierdo, A., Belcher, A. M., Scott, L., Cazares, V. A., Chen, J., O'Dell, S. J., Malvaez, M., Wu, T., & Marshall, J. F. (2010). Reversal-specific learning impairments after a binge regimen of methamphetamine in rats: possible involvement of striatal dopamine. *Neuropsychopharmacology*, *35*(2), 505-514. https://doi.org/10.1038/npp.2009.155

Izquierdo, A., Wiedholz, L. M., Millstein, R. A., Yang, R. J., Bussey, T. J., Saksida, L. M., & Holmes, A. (2006). Genetic and dopaminergic modulation of reversal learning in a touchscreen-based operant procedure for mice. *Behav Brain Res*, *171*(2), 181-188. https://doi.org/10.1016/j.bbr.2006.03.029

Jacobs, E., & D'Esposito, M. (2011). Estrogen shapes dopamine-dependent cognitive processes: implications for women's health. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *31*(14), 5286–5293. https://doi.org/10.1523/JNEUROSCI.6394-10.2011

Janak, P. H., & Tye, K. M. (2015). From circuits to behaviour in the amygdala. *Nature*, *517*(7534), 284-292. https://doi.org/10.1038/nature14188

Jang, A. I., Costa, V. D., Rudebeck, P. H., Chudasama, Y., Murray, E. A., & Averbeck, B. B. (2015). The Role of Frontal Cortical and Medial-Temporal Lobe Brain Areas in Learning a Bayesian Prior Belief on Reversals. *J Neurosci*, *35*(33), 11751-11760. https://doi.org/10.1523/JNEUROSCI.1594-15.2015

Jenni, N. L., G. Rutledge and S. B. Floresco (2022). Distinct Medial Orbitofrontal-Striatal Circuits Support Dissociable Component Processes of Risk/Reward Decision-Making. *J Neurosci* **42**(13): 2743-2755.

Jentsch, J. D., Ashenhurst, J. R., Cervantes, M. C., Groman, S. M., James, A. S., & Pennington, Z. T. (2014). Dissecting impulsivity and its relationships to drug addictions. *Annals of the New York Academy of Sciences*, *1327*(1), 1–26. https://doi.org/10.1111/nyas.12388

Jones, B., & Mishkin, M. (1972). Limbic lesions and the problem of stimulus--reinforcement associations. *Exp Neurol*, *36*(2), 362-377. https://doi.org/10.1016/0014-4886(72)90030-1

Kaluve, A. M., Le, J. T., & Graham, B. M. (2022). Female rodents are not more variable than male rodents: A meta-analysis of preclinical studies of fear and anxiety. *Neuroscience and biobehavioral reviews*, *143*, 104962. https://doi.org/10.1016/j.neubiorev.2022.104962

Karanian, J., Yergey, J., Lister, R., 'Souza, N., Linnoila, M., & Salem Jr., N. (1986). Characterization of an Automated Apparatus for Precise Control of Inhalation Chamber Ethanol Vapor and Blood Ethanol Concentrations. *Alcohol: Clinical and Experimental Research*, *10*(4), 443-447. https://doi.org/https://doi.org/10.1111/j.1530-0277.1986.tb05121.x

Keefer, S. E. and G. D. Petrovich (2022). Necessity and recruitment of cue-specific neuronal ensembles within the basolateral amygdala during appetitive reversal learning. *Neurobiology of Learning and Memory* **194**: 107663.

Kim, J., & Ragozzino, M. E. (2005). The involvement of the orbitofrontal cortex in learning under changing task contingencies. *Neurobiology of Learning and Memory*, *83*(2), 125–133. https://doi.org/10.1016/j.nlm.2004.10.003

Kimbrough, A., Kim, S., Cole, M., Brennan, M., & George, O. (2017). Intermittent Access to Ethanol Drinking Facilitates the Transition to Excessive Drinking After Chronic

Intermittent Ethanol Vapor Exposure. *Alcohol Clin Exp Res*, *41*(8), 1502-1509. https://doi.org/10.1111/acer.13434

Klenowski, P. M., Fogarty, M. J., Belmer, A., Noakes, P. G., Bellingham, M. C., & Bartlett, S. E. (2015). Structural and functional characterization of dendritic arbors and GABAergic synaptic inputs on interneurons and principal cells in the rat basolateral amygdala. *Journal of Neurophysiology*, *114*(2), 942–957. https://doi.org/10.1152/jn.00824.2014

Knapp DJ, Breese GR. (2012). Models of Chronic Alcohol Exposure and Dependence. In: Kobeissy FH, editor. *Psychiatric Disorders: Methods and Protocols* [Internet]. Totowa, NJ: Humana Press; p. 205–30. Available from: https://doi.org/10.1007/978-1-61779-458-2_13

Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: A neurocircuitry analysis. *The Lancet Psychiatry*, *3*(8), 760–773. https://doi.org/10.1016/S2215-0366(16)00104-8

Koob, G. F. (2014). Neurocircuitry of alcohol addiction: synthesis from animal models. *Handb Clin Neurol*, *125*, 33-54. https://doi.org/10.1016/b978-0-444-62619-6.00003-3

Koob, G. F., & Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *35*(1), 217–238. PubMed. https://doi.org/10.1038/npp.2009.110

Koob, G. F. (2008). A role for brain stress systems in addiction. *Neuron*, *59*(1), 11-34. https://doi.org/10.1016/j.neuron.2008.06.012

Koob, G. F. (2003). Alcoholism: Allostasis and Beyond. *Alcoholism: Clinical and Experimental Research*, *27*(2), 232–243. https://doi.org/10.1097/01.ALC.0000057122.36127.C2

Koob, G. F., & Moal, M. L. (1997). Drug Abuse: Hedonic Homeostatic Dysregulation. *Science*, *278*(5335), 52-58. https://doi.org/doi:10.1126/science.278.5335.52

Kroener, S., Mulholland, P. J., New, N. N., Gass, J. T., Becker, H. C., & Chandler, L. J. (2012). Chronic alcohol exposure alters behavioral and synaptic plasticity of the rodent prefrontal cortex. *PloS One*, *7*(5), e37541–e37541. PubMed. https://doi.org/10.1371/journal.pone.0037541

Kruse, L. C., Schindler, A. G., Williams, R. G., Weber, S. J., & Clark, J. J. (2017). Maladaptive Decision Making in Adults with a History of Adolescent Alcohol use, in a Preclinical Model, Is Attributable to the Compromised Assignment of Incentive Value during Stimulus-Reward Learning. *Frontiers in Behavioral Neuroscience*, *11*, 134. https://doi.org/10.3389/fnbeh.2017.00134

Kuzmin A, Liljequist S, Meis J, Chefer V, Shippenberg T, Bakalkin G. (2012). Repeated moderate-dose ethanol bouts impair cognitive function in Wistar rats. *Addiction Biology*. 7(1):132–40.

LaClair M, Febo M, Nephew B, Gervais NJ, Poirier G, Workman K. (2019). Sex Differences in Cognitive Flexibility and Resting Brain Networks in Middle-Aged Marmosets. *eNeuro.* 6(4):ENEURO.0154-19.2019.

Lalumiere, R. (2014). Optogenetic dissection of amygdala functioning. *Frontiers in Behavioral Neuroscience*, *8*, 107. https://doi.org/10.3389/fnbeh.2014.00107

Laughlin, R. E., Grant, T. L., Williams, R. W., & Jentsch, J. D. (2011). Genetic Dissection of Behavioral Flexibility: Reversal Learning in Mice. *Nucleus Accumbens Neuroadaptations and Relapse in Addiction*, *69*(11), 1109–1116. https://doi.org/10.1016/j.biopsych.2011.01.014

Lee, D., Seo, H., & Jung, M. W. (2012). Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci, 35*, 287-308. doi:10.1146/annurev-neuro-062111-150512

Leeman, R. F., Heilig, M., Cunningham, C. L., Stephens, D. N., Duka, T., & O'Malley, S. S. (2010). Ethanol consumption: How should we measure it? Achieving consilience between human and animal phenotypes. *Addiction Biology*, *15*(2), 109–124. https://doi.org/10.1111/j.1369-1600.2009.00192.x

Leeson, V. C., Robbins, T. W., Matheson, E., Hutton, S. B., Ron, M. A., Barnes, T. R., & Joyce, E. M. (2009). Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: stability over six years and specific associations with medication type and disorganization syndrome. *Biological psychiatry*, *66*(6), 586-593. https://doi.org/10.1016/j.biopsych.2009.05.016

Lehnert, L., & Littman, M. (2019). Successor Features Support Model-based and Model-free Reinforcement Learning. *ArXiv*, *abs/1901.11437*.

Lichtenberg, N. T., Pennington, Z. T., Holley, S. M., Greenfield, V. Y., Cepeda, C., Levine, M. S., & Wassum, K. M. (2017). Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-Triggered Reward Expectations. *The Journal of Neuroscience*, *37*(35), 8374. https://doi.org/10.1523/JNEUROSCI.0486-17.2017

Loi, B., Lobina, C., Maccioni, P., Fantini, N., Carai, M. A. M., Gessa, G. L., & Colombo, G. (2010). Increase in Alcohol Intake, Reduced Flexibility of Alcohol Drinking, and Evidence of Signs of Alcohol Intoxication in Sardinian Alcohol-Preferring Rats Exposed to Intermittent Access to 20% Alcohol. *Alcoholism: Clinical and Experimental Research*, *34*(12), 2147–2154. https://doi.org/10.1111/j.1530-0277.2010.01311.x

Loos, M., Staal, J., Smit, A., De Vries, T., & Spijker, S. (2013). Enhanced alcohol self-administration and reinstatement in a highly impulsive, inattentive recombinant inbred

mouse strain. *Frontiers in Behavioral Neuroscience*, *7*, 151. https://doi.org/10.3389/fnbeh.2013.00151

Lourdes de la Torre, M., Dolores Escarabajal, M., & Agüero, Á. (2015). Sex differences in adult Wistar rats in the voluntary consumption of ethanol after pre-exposure to ethanol-induced flavor avoidance learning. *Pharmacology Biochemistry and Behavior*, *137*, 7–15. https://doi.org/10.1016/j.pbb.2015.07.011

Louth EL, Bignell W, Taylor CL, Bailey CDC. (2016). Developmental Ethanol Exposure Leads to Long-Term Deficits in Attention and Its Underlying Prefrontal Circuitry. *eneuro.* 3(5):ENEURO.0267-16.2016.

Lovick, T. A., & Zangrossi, H., Jr (2021). Effect of Estrous Cycle on Behavior of Females in Rodent Tests of Anxiety. *Frontiers in psychiatry*, *12*, 711065. https://doi.org/10.3389/fpsyt.2021.711065

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

Lüthi, A., & Lüscher, C. (2014). Pathological circuit function underlying addiction and anxiety disorders. *Nature Neuroscience*, *17*(12), 1635–1643. https://doi.org/10.1038/nn.3849

Macoveanu, J., S. Henningsson, A. Pinborg, P. Jensen, G. M. Knudsen, V. G. Frokjaer and H. R. Siebner (2016). Sex-Steroid Hormone Manipulation Reduces Brain Response to Reward *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **41**(4): 1057-1065.

Malvaez, M., Shieh, C., Murphy, M. D., Greenfield, V. Y., & Wassum, K. M. (2019). Distinct cortical–amygdala projections drive reward value encoding and retrieval. *Nature Neuroscience*, *22*(5), 762–769. https://doi.org/10.1038/s41593-019-0374-7

Malvaez, M., Greenfield, V. Y., Wang, A. S., Yorita, A. M., Feng, L., Linker, K. E., Monbouquette, H. G., & Wassum, K. M. (2015). Basolateral amygdala rapid glutamate release encodes an outcome-specific representation vital for reward-predictive cues to selectively invigorate reward-seeking actions. *Scientific Reports*, *5*(1), 12511. https://doi.org/10.1038/srep12511

Mamlouk, G. M., Dorris, D. M., Barrett, L. R., & Meitzen, J. (2020). Sex bias and omission in neuroscience research is influenced by research model and journal, but not reported NIH funding. *Frontiers in neuroendocrinology*, *57*, 100835. https://doi.org/10.1016/j.yfrne.2020.100835

Manion, M. T. C., Glasper, E. R., & Wang, K. H. (2022). A sex difference in mouse dopaminergic projections from the midbrain to basolateral amygdala. *Biology of sex differences*, *13*(1), 75. https://doi.org/10.1186/s13293-022-00486-4

Mar, R. A. (2011). The Neural Bases of Social Cognition and Story Comprehension. *Annual Review of Psychology*, *62*(1), 103–134. https://doi.org/10.1146/annurev-psych-120709-145406

Marbach F, Zador AM. (2017). A self-initiated two-alternative forced choice paradigm for head-fixed mice. *bioRxiv.* ;073783.

Maren, S., & Fanselow, M. S. (1996). The Amygdala and Fear Conditioning: Has the Nut Been Cracked? *Neuron*, *16*(2), 237–240. https://doi.org/10.1016/S0896-6273(00)80041-0

Marquardt K, Sigdel R, Brigman JL.(2017). Touch-screen visual reversal learning is mediated by value encoding and signal propagation in the orbitofrontal cortex. N*eurobiology of learning and memory.* 139:179–88.

Martin TJ, Grigg A, Kim SA, Ririe DG, Eisenach JC. (2015). Assessment of attention threshold in rats by titration of visual cue duration during the five choice serial reaction time task. *Journal of Neuroscience Methods*. 241:37–43.

Massi, B., Donahue, C. H., & Lee, D. (2018). Volatility Facilitates Value Updating in the Prefrontal Cortex. *Neuron*, *99*(3), 598-608 e594. https://doi.org/10.1016/j.neuron.2018.06.033

MATLAB and Statistics Toolbox. Natick, Massachusetts, United States: The Mathworks, Inc.; 2013.

McCarthy, M. M., A. P. Arnold, G. F. Ball, J. D. Blaustein and G. J. De Vries (2012). Sex differences in the brain: the not so inconvenient truth. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32**(7): 2241-2247.

McDonald, A. J. (1998). Cortical pathways to the mammalian amygdala. *Progress in Neurobiology*, *55*(3), 257–332. https://doi.org/10.1016/S0301-0082(98)00003-3

McDonald, A. J., Mascagni, F., & Guo, L. (1996). Projections of the medial and lateral prefrontal cortices to the amygdala: A Phaseolus vulgaris leucoagglutinin study in the rat. *Neuroscience*, *71*(1), 55–75. https://doi.org/10.1016/0306-4522(95)00417-3

McDonald, A. J. (1992). Projection neurons of the basolateral amygdala: A correlative Golgi and retrograde tract tracing study. *Brain Research Bulletin*, *28*(2), 179–185. https://doi.org/10.1016/0361-9230(92)90177-Y

McDonald, A. J. (1991). Topographical organization of amygdaloid projections to the caudatoputamen, nucleus accumbens, and related striatal-like areas of the rat brain. *Neuroscience*, *44*(1), 15–33. https://doi.org/10.1016/0306-4522(91)90248-M

McDonald, A. J. (1982). Neurons of the lateral and basolateral amygdaloid nuclei: A golgi study in the rat. *Journal of Comparative Neurology*, *212*(3), 293–312. https://doi.org/10.1002/cne.902120307

McMurray, M. S., Amodeo, L. R., & Roitman, J. D. (2016). Consequences of Adolescent Ethanol Consumption on Risk Preference and Orbitofrontal Cortex Encoding of Reward. *Neuropsychopharmacology*, *41*(5), 1366–1375. PMC. https://doi.org/10.1038/npp.2015.288

McNamara, J. M., Fawcett, T. W., & Houston, A. I. (2013). An adaptive response to uncertainty generates positive and negative contrast effects. *Science*, *340*(6136), 1084-1086. https://doi.org/10.1126/science.1230599

Melendez, R. I. (2011). Intermittent (every-other-day) drinking induces rapid escalation of ethanol intake and preference in adolescent and adult C57BL/6J mice. *Alcoholism, Clinical and Experimental Research*, *35*(4), 652–658. PubMed. https://doi.org/10.1111/j.1530-0277.2010.01383.x

Melugin, P. R., Nolan, S. O., & Siciliano, C. A. (2021). Chapter Eight—Bidirectional causality between addiction and cognitive deficits. In E. S. Calipari & N. W. Gilpin (Eds.), *International Review of Neurobiology* (Vol. 157, pp. 371–407). Academic Press. https://doi.org/10.1016/bs.irn.2020.11.001

Meziane, H., Ouagazzal, A. M., Aubert, L., Wietrzych, M., & Krezel, W. (2007). Estrous

cycle effects on behavior of C57BL/6J and BALB/cByJ female mice: implications for phenotyping strategies. *Genes, brain, and behavior*, *6*(2), 192–200. https://doi.org/10.1111/j.1601-183X.2006.00249.x

Miller, K. M., Risher, M.-L., Acheson, S. K., Darlow, M., Sexton, H. G., Schramm-Sapyta, N., & Swartzwelder, H. S. (2017). Behavioral Inefficiency on a Risky Decision-Making Task in Adulthood after Adolescent Intermittent Ethanol Exposure in Rats. *Scientific Reports*, *7*(1), 4680. https://doi.org/10.1038/s41598-017-04704-7

Moorman, D. E., & Aston-Jones, G. (2014). Orbitofrontal Cortical Neurons Encode Expectation-Driven Initiation of Reward-Seeking. *The Journal of Neuroscience*, *34*(31), 10234. https://doi.org/10.1523/JNEUROSCI.3216-13.2014

Morales M, McGinnis MM, McCool BA. (2015). Chronic ethanol exposure increases voluntary home cage intake in adult male, but not female, Long–Evans rats. *Pharmacology Biochemistry and Behavior.* 139:67–76.

Mankes RF, Glick SD, Vanderhoeven T, Lefevre R (1991) Alcohol preference and hepatic alcohol-dehydrogenase activity in adult long-evans rats is affected by intrauterine sibling contiguity. *Alcohol Clin Exp Res* 15:80–85.

Marinelli PW, Quirion R, Gianoulakis C (2003) Estradiol valerate and alcohol intake: a comparison between wistar and lewis rats and the putative role of endorphins. Behav *Brain Res* 139:59–67.

Morrison, S. E., & Salzman, C. D. (2010). Re-valuing the amygdala. *Cognitive Neuroscience*, *20*(2), 221–230. https://doi.org/10.1016/j.conb.2010.02.007

Murray, E. A., & Izquierdo, A. (2007). Orbitofrontal Cortex and Amygdala Contributions to Affect and Action in Primates. *Annals of the New York Academy of Sciences*, *1121*(1), 273–296. https://doi.org/10.1196/annals.1401.021

Namboodiri, K. V. M., T. Hobbs, I. Trujillo-Pisanty, R. C. Simon, M. M. Gray and G. D. Stuber (2021). Relative salience signaling within a thalamo-orbitofrontal circuit governs learning rate. *Current biology : CB* **31**(23): 5176-5191.e5175.

Namboodiri, V. M. K., J. M. Otis, K. van Heeswijk, E. S. Voets, R. A. Alghorazi, J. Rodriguez-Romaguera, S. Mihalas and G. D. Stuber (2019). Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. *Nat Neurosci* **22**(7): 1110-1121.

Nieto, S. J., Grodin, E. N., Aguirre, C. G., Izquierdo, A., & Ray, L. A. (2021). Translational opportunities in animal and human models to study alcohol use disorder. *Transl Psychiatry*, *11*(1), 496. https://doi.org/10.1038/s41398-021-01615-0

NIH. 2015. Consideration of sex as a biological variable in NIH-funded research. National Institutes of HTealth Notice Number: NOT-OD-15–102. National Institutes of Health. https://grants.nih.gov/grants/guide/noticefiles/not-od-15-102.html

Nixon K, Kim DH, Potts EN, He J, Crews FT. (2008). Distinct cell proliferation events during abstinence after alcohol dependence: Microglia proliferation precedes neurogenesis. *Neurobiology of Disease.* 31(2):218–29.

Nixon K, Crews FT. (2004). Temporally Specific Burst in Cell Proliferation Increases Hippocampal Neurogenesis in Protracted Abstinence from Alcohol. *J Neurosci.* 24(43):9714.

Noonan, M. P., Kolling, N., Walton, M. E., & Rushworth, M. F. S. (2012). Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *European Journal of Neuroscience*, *35*(7), 997–1010. https://doi.org/10.1111/j.1460-9568.2012.08023.x

O'Neill J, Cardenas VA, Meyerhoff DJ. (2001). Effects of Abstinence on the Brain: Quantitative Magnetic Resonance Imaging and Magnetic Resonance Spectroscopic Imaging in Chronic Alcohol Abuse. *Alcoholism: Clinical and Experimental Research.* 25(11):1673–82.

Ongür, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral cortex (New York, N.Y. : 1991)*, *10*(3), 206-219. https://doi.org/10.1093/cercor/10.3.206

Orsini, C. A., L. M. Truckenbrod and A.-R. Wheeler (2022). Regulation of sex differences in risk-based decision making by gonadal hormones: Insights from rodent models. *Behavioural processes* **200**: 104663-104663.

Orsini, C. A. and B. Setlow (2017). Sex differences in animal models of decision making. *Journal of neuroscience research* **95**(1-2): 260-269.

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 neuroscienc*e, 130(1), 50–61. https://doi.org/10.1037/bne0000111

Oscar-Berman, M., Kirkley, S. M., Gansler, D. A., & Couture, A. (2004). Comparisons of Korsakoff and non-Korsakoff alcoholics on neuropsychological tests of prefrontal brain functioning. *Alcoholism, Clinical and Experimental Research*, *28*(4), 667–675. PubMed.

Ostlund, S. B., & Balleine, B. W. (2008). Differential Involvement of the Basolateral Amygdala and Mediodorsal Thalamus in Instrumental Action Selection. *The Journal of Neuroscience*, *28*(17), 4398. https://doi.org/10.1523/JNEUROSCI.5472-07.2008

Ostlund, S. B., & Balleine, B. W. (2007a). Orbitofrontal Cortex Mediates Outcome Encoding in Pavlovian But Not Instrumental Conditioning. *The Journal of Neuroscience*, *27*(18), 4819. https://doi.org/10.1523/JNEUROSCI.5443-06.2007

Ostlund, S. B., & Balleine, B. W. (2007b). The Contribution of Orbitofrontal Cortex to Action Selection. *Annals of the New York Academy of Sciences*, *1121*(1), 174–192. https://doi.org/10.1196/annals.1401.033

Ottersen, O. P. (1982). Connections of the amygdala of the rat. IV: Corticoamygdaloid and intraamygdaloid connections as studied with axonal transport of horseradish peroxidase. *Journal of Comparative Neurology*, *205*(1), 30–48. https://doi.org/10.1002/cne.902050104

Overman WH. (2004). Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. *Brain and Cognition.* 55(1):134–47.

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L. F., Dalgleish, H., Carandini, M., & Harris, K. D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv*, 061507. https://doi.org/10.1101/061507

Padoa-Schioppa, C., & Conen, K. E. (2017). Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron*, *96*(4), 736–754. PubMed. https://doi.org/10.1016/j.neuron.2017.09.031

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. PubMed. https://doi.org/10.1038/nature04676

Palanza, P., Gioiosa, L., & Parmigiani, S. (2001). Social stress in mice: gender differences and effects of estrous cycle and social dominance. *Physiol Behav*, *73*(3), 411-420. https://doi.org/10.1016/s0031-9384(01)00494-2

Pamplona-Santos, D., Lamarão-Vieira, K., Nascimento, P. C., Bittencourt, L. O., Corrêa, M. G., Dos Santos, S. M., Cartágenes, S. C., Fernandes, L. M. P., Monteiro, M. C., Maia, C. S. F., & Lima, R. R. (2019). Aerobic Physical Exercise as a Neuroprotector Strategy for Ethanol Binge-Drinking Effects in the Hippocampus and Systemic Redox Status in Rats. *Oxidative medicine and cellular longevity*, *2019*, 2415243. https://doi.org/10.1155/2019/2415243

Panayi, M. C., & Killcross, S. (2018). Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. *ELife*, *7*, e37357. https://doi.org/10.7554/eLife.37357

Parkes, S., Ravassard, P., Cerpa, J.-C., Wolff, M., Ferreira, G., & Coutureau, E. (2017). Insular and Ventrolateral Orbitofrontal Cortices Differentially Contribute to Goal-Directed Behavior in Rodents. *Cerebral Cortex (New York, N.Y. : 1991)*, *28*, 1–13. https://doi.org/10.1093/cercor/bhx132

Paton, J. J., Belova, M. A., Morrison, S. E., & Salzman, C. D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*, *439*(7078), 865–870. https://doi.org/10.1038/nature04490

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

Pellman, B. A., Schuessler, B. P., Tellakat, M., & Kim, J. J. (2017). Sexually Dimorphic Risk Mitigation Strategies in Rats. *eNeuro*, *4*(1), ENEURO.0288-16.2017. https://doi.org/10.1523/ENEURO.0288-16.2017

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior. *Neuron*, *48*(2), 175–187. https://doi.org/10.1016/j.neuron.2005.09.025

Piano MR, Carrigan TM, Schwertz DW. (2005). Sex differences in ethanol liquid diet consumption in Sprague–Dawley rats. *Alcohol.* 35(2):113–8.

Piantadosi P, Lieberman A, Pickens C, Bergstrom H, Holmes A. (2019). A novel multichoice touchscreen paradigm for assessing cognitive flexibility in mice. *Learning & Memory.* 26:24–30.

Pickens, C. L., Saddoris, M. P., Gallagher, M., & Holland, P. C. (2005). Orbitofrontal Lesions Impair Use of Cue-Outcome Associations in a Devaluation Task. *Behavioral Neuroscience*, *119*(1), 317–322. https://doi.org/10.1037/0735-7044.119.1.317

Pickens, C. L., Saddoris, M. P., Setlow, B., Gallagher, M., Holland, P. C., & Schoenbaum, G. (2003). Different Roles for Orbitofrontal Cortex and Basolateral Amygdala in a Reinforcer Devaluation Task. *The Journal of Neuroscience*, *23*(35), 11078. https://doi.org/10.1523/JNEUROSCI.23-35-11078.2003

Pitel, A. L., Rivier, J., Beaunieux, H., Vabret, F., Desgranges, B., & Eustache, F. (2009). Changes in the Episodic Memory and Executive Functions of Abstinent and Relapsed Alcoholics Over a 6-Month Period. *Alcoholism: Clinical and Experimental Research*, *33*(3), 490–498. https://doi.org/10.1111/j.1530-0277.2008.00859.x

Planeta, C. S. (2013). Animal models of alcohol and drug dependence. *Revista Brasileira de Psiquiatria*, *35*, S140–S146.

Prendergast, B. J., Onishi, K. G., & Zucker, I. (2014). Female mice liberated for

inclusion in neuroscience and biomedical research. *Neuroscience and biobehavioral reviews*, *40*, 1–5. https://doi.org/10.1016/j.neubiorev.2014.01.001

Priddy BM, Carmack SA, Thomas LC, Vendruscolo JCM, Koob GF, Vendruscolo LF. (2017). Sex, strain, and estrous cycle influences on alcohol drinking in rats. *Pharmacology Biochemistry and Behavior.* 152:61–7.

Quirarte GL, Reid LD, de la Teja IS, Reid ML, Sanchez MA, Diaz-Trujillo A, Aguilar-Vazquez A, Prado-Alcala RA (2007) Estradiol valerate and alcohol intake: dose-response assessments. *BMC Pharmacol* 7:3.

Ragozzino, M. E. (2007). The Contribution of the Medial Prefrontal Cortex, Orbitofrontal Cortex, and Dorsomedial Striatum to Behavioral Flexibility. *Annals of the New York Academy of Sciences*, *1121*(1), 355–375. https://doi.org/10.1196/annals.1401.013

Rajasingh J, Bord E, Qin G, Ii M, Silver M, Hamada H, AhluwaliaD, Goukassian D, Zhu Y, Losordo DW, Kishore R (2007)Enhanced voluntary alcohol consumption after estrogen sup-plementation negates estrogen-mediated vascular repair inovariectomized mice. *Endocrinology* 148:3618–3624.

Reid ML, Hubbell CL, Reid LD (2003) A pharmacological dose ofestradiol can enhance appetites for alcoholic beverages.*Pharmacol Biochem Behav* 74:381–388.

Reid LD,Marinelli PW, Bennett SM, Fiscale LT, Narciso SP,Oparowski CJ, Reid ML, Merrigan BA, Moricone J, HubbellCL, Gianoulakis C (2002) One injection of estradiol valerateinduces dramatic changes in rats'intake of alcoholic bever-ages. *Pharmacol Biochem Behav* 72:601–616.

Ray MH, Hite T, Gallo M, Pickens CL. (2018) Operant over-responding is more sensitive than reversal learning for revealing behavioral changes after withdrawal from alcohol consumption. *Physiology & Behavior.* 1;196:176–84.

Ray, J. P., & Price, J. L. (1992). The organization of the thalamocortical connections of the mediodorsal thalamic nucleus in the rat, related to the ventral forebrain–prefrontal cortex topography. *Journal of Comparative Neurology*, *323*(2), 167–197. https://doi.org/10.1002/cne.903230204

Remijnse, P. L., Nielen, M. M., van Balkom, A. J., Cath, D. C., van Oppen, P., Uylings, H. B., & Veltman, D. J. (2006). Reduced orbitofrontal-striatal activity on a reversal learning task in obsessive-compulsive disorder. *Arch Gen Psychiatry*, *63*(11), 1225-1236. https://doi.org/10.1001/archpsyc.63.11.1225

Riceberg, J. S., & Shapiro, M. L. (2017). Orbitofrontal Cortex Signals Expected Outcomes with Predictive Codes When Stable Contingencies Promote the Integration of Reward History. *The Journal of Neuroscience*, *37*(8), 2010. https://doi.org/10.1523/JNEUROSCI.2951-16.2016

Riceberg, J. S., & Shapiro, M. L. (2012). Reward Stability Determines the Contribution of Orbitofrontal Cortex to Adaptive Behavior. *The Journal of Neuroscience*, *32*(46), 16402. https://doi.org/10.1523/JNEUROSCI.0776-12.2012

Robbins, T. W., & Cardinal, R. N. (2019). Computational psychopharmacology: A translational and pragmatic approach. *Psychopharmacology*, *236*(8), 2295–2305. https://doi.org/10.1007/s00213-019-05302-3

Robbins, T. W. (2002). The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. *Psychopharmacology (Berl)*, *163*(3-4), 362-380. https://doi.org/10.1007/s00213-002-1154-7

Robinson, E. S., Eagle, D. M., Economidou, D., Theobald, D. E., Mar, A. C., Murphy, E. R., Robbins, T. W., & Dalley, J. W. (2009). Behavioural characterisation of high impulsivity on the 5-choice serial reaction time task: specific deficits in 'waiting' versus 'stopping'. *Behav Brain Res*, *196*(2), 310-316. https://doi.org/10.1016/j.bbr.2008.09.021

Rodd-Henricks, Z. A., McKinzie, D. L., Murphy, J. M., McBride, W. J., Lumeng, L., & Li, T.-K. (2000). The Expression of an Alcohol Deprivation Effect in the High–Alcohol-Drinking Replicate Rat Lines Is Dependent On Repeated Deprivations. *Alcoholism: Clinical and Experimental Research*, *24*(6), 747–753. https://doi.org/10.1111/j.1530-0277.2000.tb02051.x

Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. *Eur J Neurosci*, *35*(7), 1190-1200. https://doi.org/10.1111/j.1460-9568.2011.07986.x

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

Roesch, M. R., G. R. Esber, D. W. Bryden, D. H. Cerri, Z. R. Haney and G. Schoenbaum (2012). Normal aging alters learning and attention-related teaching signals in basolateral amygdala. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32**(38): 13137-13144.

Rosenbloom MJ, Rohlfing T, O'Reilly AW, Sassoon SA, Pfefferbaum A, Sullivan EV. (2007). Improvement in memory and static balance with abstinence in alcoholic men and women: Selective relations with change in brain structure. *Psychiatry Research: Neuroimaging.* 155(2):91–102.

Rosenwasser AM, Fixaris MC, Crabbe JC, Brooks PC, Ascheid S. (2012). Escalation of intake under intermittent ethanol access in diverse mouse genotypes. *Addiction Biology*. 18(3):496–507.

Rudebeck, P.H. (2013). Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **33**(9): 4105-4109.

Rudebeck, P. H., & Murray, E. A. (2008). Amygdala and Orbitofrontal Cortex Lesions Differentially Influence Choices during Object Reversal Learning. *The Journal of Neuroscience*, *28*(33), 8338. https://doi.org/10.1523/JNEUROSCI.2272-08.2008

Saddoris, M. P., Gallagher, M., & Schoenbaum, G. (2005). Rapid Associative Encoding in Basolateral Amygdala Depends on Connections with Orbitofrontal Cortex. *Neuron*, *46*(2), 321–331. https://doi.org/10.1016/j.neuron.2005.02.018

Salvatore, M., Wiersielis, K. R., Luz, S., Waxler, D. E., Bhatnagar, S., & Bangasser, D. A. (2018). Sex differences in circuits activated by corticotropin releasing factor in rats. *Hormones and behavior*, *97*, 145–153. https://doi.org/10.1016/j.yhbeh.2017.10.004

Satta, R., Hilderbrand, E. R., & Lasek, A. W. (2018). Ovarian Hormones Contribute to High Levels of Binge-Like Drinking by Female Mice. *Alcoholism, clinical and experimental research*, *42*(2), 286–294. https://doi.org/10.1111/acer.13571

Sanchez-Roige S, Baro V, Trick L, Peña-Oliver Y, Stephens DN, Duka T. (2014). Exaggerated Waiting Impulsivity Associated with Human Binge Drinking, and High Alcohol Consumption in Mice. *Neuropsychopharmacology.* 39(13):2919–27.

Sanchez-Roige S, Peña-Oliver Y, Ripley TL, Stephens DN. (2014). Repeated Ethanol Exposure During Early and Late Adolescence: Double Dissociation of Effects on Waiting and Choice Impulsivity. *Alcoholism: Clinical and Experimental Research.* 38(10):2579–89.

Schoenbaum, G., Setlow, B., Saddoris, M. P., & Gallagher, M. (2006). Encoding Changes in Orbitofrontal Cortex in Reversal-Impaired Aged Rats. *Journal of Neurophysiology*, *95*(3), 1509–1517. https://doi.org/10.1152/jn.01052.2005

Schoenbaum, G., Setlow, B., Nugent, S. L., Saddoris, M. P., & Gallagher, M. (2003). Lesions of Orbitofrontal Cortex and Basolateral Amygdala Complex Disrupt Acquisition of Odor-Guided Discriminations and Reversals. *Learning & Memory*, *10*(2), 129–140. https://doi.org/10.1101/lm.55203

Schoenbaum, G., Nugent, S. L., Saddoris, M. P., & Setlow, B. (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *NeuroReport*, *13*(6).https://journals.lww.com/neuroreport/Fulltext/2002/05070/Orbitofrontal_lesions_in_rats_impair_reversal_but.30.aspx

Schoenbaum, G., Chiba, A. A., & Gallagher, M. (2000). Changes in Functional Connectivity in Orbitofrontal Cortex and Basolateral Amygdala during Learning and Reversal Training. *The Journal of Neuroscience*, *20*(13), 5179. https://doi.org/10.1523/JNEUROSCI.20-13-05179.2000

Schoenbaum, G., Chiba, A. A., & Gallagher, M. (1999). Neural Encoding in Orbitofrontal Cortex and Basolateral Amygdala during Olfactory Discrimination Learning. *The Journal of Neuroscience*, *19*(5), 1876. https://doi.org/10.1523/JNEUROSCI.19-05-01876.1999

Schramm-Sapyta NL, Francis R, MacDonald A, Keistler C, O'Neill L, Kuhn CM. (2014). Effect of sex on ethanol consumption and conditioned taste aversion in adolescent and adult rats. *Psychopharmacology.* 231(8):1831–9.

Shansky, R. M. and A. Z. Murphy (2021). Considering sex as a biological variable will require a global shift in science culture. *Nature Neuroscience* **24**(4): 457-464.

Shansky, R. M., Hamo, C., Hof, P. R., McEwen, B. S., & Morrison, J. H. (2009). Stress-induced dendritic remodeling in the prefrontal cortex is circuit specific. *Cerebral cortex (New York, N.Y. : 1991)*, *19*(10), 2479–2484. https://doi.org/10.1093/cercor/bhp003

Sherrill LK, Berthold C, Koss WA, Juraska JM, Gulley JM. (2011). Sex differences in the effects of ethanol pre-exposure during adolescence on ethanol-induced conditioned taste aversion in adult rats. *Behavioural Brain Research.* 225(1):104–9.

Sherrill LK, Koss WA, Foreman ES, Gulley JM (2011) The effects of pre-pubertal gonadectomy and binge-like ethanol exposure during adolescence on ethanol drinking in adult male and female rats. *Behav Brain Res* 216:569–575.

Shnitko, T. A., Gonzales, S. W., & Grant, K. A. (2019). Low cognitive flexibility as a risk for heavy alcohol drinking in non-human primates. *New Technologies in Alcohol Research and Treatment*, *74*, 95–104. https://doi.org/10.1016/j.alcohol.2018.04.007

Sias, A. C., Morse, A. K., Wang, S., Greenfield, V. Y., Goodpaster, C. M., Wrenn, T. M., Wikenheiser, A. M., Holley, S. M., Cepeda, C., Levine, M. S., & Wassum, K. M. (2021).

A bidirectional corticoamygdala circuit for the encoding and retrieval of detailed reward memories. *ELife*, *10*, e68617. https://doi.org/10.7554/eLife.68617

Simms, J. A., Steensland, P., Medina, B., Abernathy, K. E., Chandler, L. J., Wise, R., & Bartlett, S. E. (2008). Intermittent Access to 20% Ethanol Induces High Ethanol Consumption in Long–Evans and Wistar Rats. *Alcoholism, Clinical and Experimental Research*, *32*(10), 1816–1823. PMC. https://doi.org/10.1111/j.1530-0277.2008.00753.x

Sinnott RS, Phillips TJ, Finn DA (2002) Alteration of voluntary ethanol and saccharin consumption by the neurosteroid allopregnanolone in mice. *Psychopharmacology* (Berl) 162:438–447.

Slawecki CJ. (2016). Two-choice reaction time performance in Sprague–Dawley rats exposed to alcohol during adolescence or adulthood. *Behavioural Pharmacology* 17(7). https://journals.lww.com/behaviouralpharm/Fulltext/2006/11000/Two_choice_reaction_time_performance_in.6.aspx

Soltani, A., & Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. *Nature Reviews Neuroscience*, *20*(10), 635–644. https://doi.org/10.1038/s41583-019-0180-y

Soltani, A., & Koechlin, E. (2021). Computational models of adaptive behavior and prefrontal cortex. *Neuropsychopharmacology*. https://doi.org/10.1038/s41386-021-01123-1

Soltani, A., Noudoost, B., & Moore, T. (2013). Dissociable dopaminergic control of saccadic target selection and its implications for reward modulation. *Proc Natl Acad Sci U S A*, *110*(9), 3579-3584. https://doi.org/10.1073/pnas.1221236110

Song, M. R., & Lee, S. W. (2020). Dynamic resource allocation during reinforcement learning accounts for ramping and phasic dopamine activity. *Neural Netw*, *126*, 95-107. https://doi.org/10.1016/j.neunet.2020.03.005

Spampanato, J., Polepalli, J., & Sah, P. (2011). Interneurons in the basolateral amygdala. *Synaptic Plasticity & Interneurons*, *60*(5), 765–773. https://doi.org/10.1016/j.neuropharm.2010.11.006

Spoelder, M., Flores Dourojeanni, J. P., de Git, K. C. G., Baars, A. M., Lesscher, H. M. B., & Vanderschuren, L. J. M. J. (2017). Individual differences in voluntary alcohol intake in rats: Relationship with impulsivity, decision making and Pavlovian conditioned approach. *Psychopharmacology*, *234*(14), 2177–2196. https://doi.org/10.1007/s00213-017-4617-6

Stalnaker, T. A., Franz, T. M., Singh, T., & Schoenbaum, G. (2007). Basolateral Amygdala Lesions Abolish Orbitofrontal-Dependent Reversal Impairments. *Neuron*, *54*(1), 51–58. https://doi.org/10.1016/j.neuron.2007.02.014

Stavro, K., Pelletier, J., & Potvin, S. (2013). Widespread and sustained cognitive deficits in alcoholism: A meta-analysis. *Addiction Biology*, *18*(2), 203–213. https://doi.org/10.1111/j.1369-1600.2011.00418.x

Steiner, A., & Redish, A. D. (2012). The Road Not Taken: Neural Correlates of Decision Making in Orbitofrontal Cortex. *Frontiers in Neuroscience*, *6*, 131. https://doi.org/10.3389/fnins.2012.00131

Stolyarova, A., Rakhshan, M., Hart, E. E., O'Dell, T. J., Peters, M. A. K., Lau, H., Soltani, A., & Izquierdo, A. (2019). Contributions of anterior cingulate cortex and basolateral amygdala

to decision confidence and learning under uncertainty. *Nature Communications*, *10*(1), 4704. https://doi.org/10.1038/s41467-019-12725-1

Stolyarova, A., & Izquierdo, A. (2017). Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. *Elife*, *6*. https://doi.org/10.7554/eLife.27483

Stolyarova, A., O'Dell, S. J., Marshall, J. F., & Izquierdo, A. (2014). Positive and negative feedback learning and associated dopamine and serotonin transporter binding after methamphetamine. *Behav Brain Res*, *271*, 195-202. https://doi.org/10.1016/j.bbr.2014.06.031

Sul, J. H., Kim, H., Huh, N., Lee, D., & Jung, M. W. (2010). Distinct Roles of Rodent Orbitofrontal and Medial Prefrontal Cortex in Decision Making. *Neuron*, *66*(3), 449–460. https://doi.org/10.1016/j.neuron.2010.03.033

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. MIT Press; /z-wcorg/.

Swainson, R., Rogers, R. D., Sahakian, B. J., Summers, B. A., Polkey, C. E., & Robbins, T. W. (2000). Probabilistic learning and reversal deficits in patients with Parkinson's disease or frontal or temporal lobe lesions: possible adverse effects of dopaminergic medication. *Neuropsychologia*, *38*(5), 596-612. https://doi.org/10.1016/s0028-3932(99)00103-7

Tada, M., Takeuchi, A., Hashizume, M., Kitamura, K., & Kano, M. (2014). A highly sensitive fluorescent indicator dye for calcium imaging of neural activity in vitro and in vivo. *Eur J Neurosci*, *39*(11), 1720-1728. https://doi.org/10.1111/ejn.12476

Taswell, C. A., Costa, V. D., Basile, B. M., Pujara, M. S., Jones, B., Manem, N., Murray, E. A., & Averbeck, B. B. (2021). Effects of Amygdala Lesions on Object-Based Versus Action-

Based Learning in Macaques. *Cereb Cortex*, *31*(1), 529-546. https://doi.org/10.1093/cercor/bhaa241

Taylor, J. R. (2019). Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes. *Neuron* **103**(4): 734-746 e733.

Torres, O. V., Walker, E. M., Beas, B. S., & O'Dell, L. E. (2014). Female rats display enhanced rewarding effects of ethanol that are hormone dependent. *Alcohol Clin Exp Res*, *38*(1), 108-115. https://doi.org/10.1111/acer.12213

Trepka, E., Spitmaan, M., Bari, B. A., Costa, V. D., Cohen, J. Y., & Soltani, A. (2021). Novel entropy-based metrics for predicting choice behavior based on local response to reward. *BioRxiv*, 2021.05.20.445009. https://doi.org/10.1101/2021.05.20.445009

Trick, L., Kempton, M. J., Williams, S. C. R., & Duka, T. (2014). Impaired fear recognition and attentional set-shifting is associated with brain structural changes in alcoholic patients. *Addiction Biology*, *19*(6), 1041-1054. https://doi.org/https://doi.org/10.1111/adb.12175

Turner KM, Peak J, Burne THJ. (2016). Measuring Attention in Rodents: Comparison of a Modified Signal Detection Task and the 5-Choice Serial Reaction Time Task. *Front Behav Neurosci.* 9:370–370.

Tye, K. M., & Janak, P. H. (2007). Amygdala neurons differentially encode motivation and reinforcement. *J Neurosci*, *27*(15), 3937-3945. https://doi.org/10.1523/JNEUROSCI.5281-06.2007

Uddin, L. Q. (2021). Cognitive and behavioural flexibility: neural mechanisms and clinical considerations. *Nature Reviews Neuroscience*, *22*(3), 167-179. https://doi.org/10.1038/s41583-021-00428-w

Uylings, H. B., & van Eden, C. G. (1990). Qualitative and quantitative comparison of the prefrontal cortex in rat and in primates, including humans. *Prog Brain Res*, *85*, 31-62. https://doi.org/10.1016/s0079-6123(08)62675-8

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 Schaaf, M. E., Zwiers, M. P., van Schouwenburg, M. R., Geurts, D. E., Schellekens, A. F., Buitelaar, J. K., Verkes, R. J., & Cools, R. (2013). Dopaminergic drug effects during reversal learning depend on anatomical connections between the orbitofrontal cortex and the amygdala. *Frontiers in neuroscience*, *7*, 142. https://doi.org/10.3389/fnins.2013.00142

van Hasselt, F. N., L. de Visser, J. M. Tieskens, S. Cornelisse, A. M. Baars, M. Lavrijsen, H. J. Krugers, R. van den Bos and M. Joëls (2012). Individual variations in maternal care early in life correlate with later life decision-making and c-fos expression in prefrontal subregions of rats. *PloS one 7*(5): e37820-e37820.

van Vulpen, E. H. S., & W.H. Verwer, R. (1989). Organization of projections from the mediodorsal nucleus of the thalamus to the basolateral complex of the amygdala in the rat. *Brain Research*, *500*(1), 389–394. https://doi.org/10.1016/0006-8993(89)90337-5

Verharen, J. P. H., den Ouden, H. E. M., Adan, R. A. H., & Vanderschuren, L. J. M. J. (2020). Modulation of value-based decision making behavior by subregions of the rat prefrontal cortex. *Psychopharmacology*. https://doi.org/10.1007/s00213-020-05454-7

Verharen, J. P. H., Kentrop, J., Vanderschuren, L. J. M. J., & Adan, R. A. H. (2019).

Reinforcement learning across the rat estrous cycle. *Psychoneuroendocrinology*, *100*, 27–31. https://doi.org/10.1016/j.psyneuen.2018.09.016

Vetreno, R. P., Bohnsack, J. P., Kusumo, H., Liu, W., Pandey, S. C., & Crews, F. T. (2020). Neuroimmune and epigenetic involvement in adolescent binge ethanol-induced loss of basal forebrain cholinergic neurons: Restoration with voluntary exercise. *Addiction biology*, *25*(2), e12731. https://doi.org/10.1111/adb.12731

Vetreno, R. P., & Crews, F. T. (2015). Binge ethanol exposure during adolescence leads to a persistent loss of neurogenesis in the dorsal and ventral hippocampus that is associated with impaired adult cognitive functioning. *Frontiers in neuroscience*, *9*, 35. https://doi.org/10.3389/fnins.2015.00035

Vetter-O'Hagen, C., Varlinskaya, E., & Spear, L. (2009). Sex Differences in Ethanol Intake and Sensitivity to Aversive Effects during Adolescence and Adulthood. *Alcohol and Alcoholism*, *44*(6), 547–554. https://doi.org/10.1093/alcalc/agp048

Vetter-O'Hagen, C. S., & Spear, L. P. (2011). The effects of gonadectomy on age- and sex-typical patterns of ethanol consumption in sprague-dawley rats. *Alcoholism, Clinical and Experimental Research*, *35*(11), 2039–2049. PMC. https://doi.org/10.1111/j.1530-0277.2011.01555.x

Wallin-Miller KG, Chesley J, Castrillon J, Wood RI. (2017). Sex differences and hormonal modulation of ethanol-enhanced risk taking in rats. *Drug and Alcohol Dependence.* 174:137–44.

Wang, A. Y., Miura, K. & Uchida, N. The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat. Neurosci*. 16, 639–647 (2013).

Wang, S.-H., Ostlund, S. B., Nader, K., & Balleine, B. W. (2005). Consolidation and Reconsolidation of Incentive Learning in the Amygdala. *The Journal of Neuroscience*, *25*(4), 830. https://doi.org/10.1523/JNEUROSCI.4716-04.2005

Ward, R. D., V. Winiger, E. R. Kandel, P. D. Balsam and E. H. Simpson (2015). Orbitofrontal cortex mediates the differential impact of signaled-reward probability on discrimination accuracy. *Front Neurosci* **9**: 230.

Warren, J. G., Fallon, V. M., Goodwin, L., Gage, S. H., & Rose, A. K. (2021). Menstrual Cycle Phase, Hormonal Contraception, and Alcohol Consumption in Premenopausal Females: A Systematic Review. *Frontiers in global women's health*, *2*, 745263. https://doi.org/10.3389/fgwh.2021.745263

Wassum, K. M., & Izquierdo, A. (2015). The basolateral amygdala in reward learning and addiction. *Neurosci Biobehav Rev*, *57*, 271-283. https://doi.org/10.1016/j.neubiorev.2015.08.017

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. PubMed. https://doi.org/10.1016/j.neuron.2013.11.005

Winstanley, C. A., & Floresco, S. B. (2016). Deciphering Decision Making: Variation in Animal Models of Effort- and Uncertainty-Based Choice Reveals Distinct Neural Circuitries Underlying Core Cognitive Processes. *J Neurosci*, *36*(48), 12069-12079. https://doi.org/10.1523/JNEUROSCI.1713-16.2016

Wise, R. A. (1973). Voluntary ethanol intake in rats following exposure to ethanol on various schedules. *Psychopharmacologia*, *29*(3), 203–210. https://doi.org/10.1007/BF00414034

Wilhelm Clare J., Hashimoto Joel G., Roberts Melissa L., Bloom Shelley H., Andrew Melissa R., Wiren Kristine M. (2015). Astrocyte Dysfunction Induced by Alcohol in Females but Not Males. *Brain Pathology*. 26(4):433–51.

Will, T. R., Proaño, S. B., Thomas, A. M., Kunz, L. M., Thompson, K. C., Ginnari, L. A., Jones, C. H., Lucas, S. C., Reavis, E. M., Dorris, D. M., & Meitzen, J. (2017). Problems and Progress regarding Sex Bias and Omission in Neuroscience Research. *eNeuro*, *4*(6), ENEURO.0278-17.2017. https://doi.org/10.1523/ENEURO.0278-17.2017

Williams, O. O. F., Coppolino, M., George, S. R., & Perreault, M. L. (2021). Sex Differences in Dopamine Receptors and Relevance to Neuropsychiatric Disorders. *Brain sciences*, *11*(9), 1199. https://doi.org/10.3390/brainsci11091199

Woitowich, N. C., Beery, A., & Woodruff, T. (2020). A 10-year follow-up study of sex inclusion in the biological sciences. *eLife*, *9*, e56344. https://doi.org/10.7554/eLife.56344

Wolff, N., Gussek, P., Stock, A.-K., & Beste, C. (2018). Effects of high-dose ethanol intoxication and hangover on cognitive flexibility. *Addiction Biology*, *23*(1), 503-514. https://doi.org/https://doi.org/10.1111/adb.12470

Worthy DA, Maddox WT. (2014). A Comparison Model of Reinforcement-Learning and Win-Stay-Lose-Shift Decision-Making Processes: A Tribute to W.K. Estes. *Journal of mathematical psychology*. 59:41–9.

Wright, S. L., G. M. Martin, C. M. Thorpe, K. Haley and D. M. Skinner (2019). Distance and direction, but not light cues, support response reversal learning. *Learning & Behavior* **47**(1): 38-46.

Xin, B., Yu, H., Qin, Y., Tang, Q., & Zhu, Z. (2020). Exploration Entropy for Reinforcement Learning. *Mathematical Problems in Engineering*, *2020*, 2672537. https://doi.org/10.1155/2020/2672537

Ye, T., J. L. Romero-Sosa, A. Rickard, C. G. Aguirre, A. M. Wikenheiser, H. T. Blair and A. Izquierdo (2021). Theta oscillations in anterior cingulate cortex and orbitofrontal cortex differentially modulate accuracy and speed in flexible reward learning. *biorxiv.org.*

Yoest, K. E., Cummings, J. A., & Becker, J. B. (2019). Oestradiol influences on dopamine release from the nucleus accumbens shell: sex differences and the role of selective oestradiol receptor subtypes. *British journal of pharmacology*, *176*(21), 4136–4148. https://doi.org/10.1111/bph.14531

Yoest, K. E., Cummings, J. A., & Becker, J. B. (2014). Estradiol, dopamine and motivation. *Central nervous system agents in medicinal chemistry*, *14*(2), 83–89. https://doi.org/10.2174/1871524914666141226103135

Zimmermann, K. S., Li, C. C., Rainnie, D. G., Ressler, K. J., & Gourley, S. L. (2018). Memory Retention Involves the Ventrolateral Orbitofrontal Cortex: Comparison with the Basolateral Amygdala. *Neuropsychopharmacology*, *43*(2), 373-383. https://doi.org/10.1038/npp.2017.139