




ARTICLE

A novel multidimensional reinforcement task in mice elucidates sex-specific behavioral strategies

Munir Gunes Kutlu¹, Jennifer E. Zachry¹, Lillian J. Brady¹, Patrick R. Melugin², Shannon J. Kelly¹, Christina Sanders¹, Jennifer Tat¹, Amy R. Johnson¹, Kimberly Thibeault², Alberto J. Lopez¹, Cody A. Siciliano^{1,2,5} and Erin S. Calipari^{1,2,3,4,5} 

A large body of work has focused on understanding stimulus-driven behavior, sex differences in these processes, and the neural circuits underlying them. Many preclinical mouse models present rewarding or aversive stimuli in isolation, ignoring that ethologically, reward seeking requires the consideration of potential aversive outcomes. In addition, the context (or reinforcement schedule under) in which stimuli are encountered can engender different behavioral responses to the same stimulus. Thus, delineating neural control of behavior requires a dissociation between stimulus valence and stimulus-driven behavior. We developed the Multidimensional Cue Outcome Action Task (MCOAT) to dissociate motivated action from cue learning and valence in mice. First, mice acquire positive and negative reinforcement in the presence of discrete discriminative stimuli. Next, discriminative stimuli are presented concurrently allowing for parsing innate behavioral strategies based on reward seeking and avoidance. Lastly, responding in the face of punishment is assessed, thus examining how positive and negative outcomes are relatively valued. First, we identified sex-specific behavioral strategies, showing that females prioritize avoidance of negative outcomes over seeking positive, while males have the opposite strategy. Next, we show that chemogenetically inhibiting D1 medium spiny neurons (MSNs) in the nucleus accumbens—a population that has been linked to reward-driven behavior—reduces positive and increases negative reinforcement learning rates. Thus, D1 MSNs modulate stimulus processing, rather than motivated responses or the reinforcement process itself. Together, the MCOAT has broad utility for understanding complex behaviors as well as the definition of the discrete information encoded within cellular populations.

Neuropsychopharmacology (2020) 45:1463–1472; <https://doi.org/10.1038/s41386-020-0692-1>

INTRODUCTION

Most of our understanding about behavioral strategies gained from preclinical research relies on unitary measures of behavior where animals have one option in the environment and often across only one reinforcer. As such, unidimensional behavioral tasks are not sufficient to gain a holistic understanding of behavioral functions. In recent years, research focused on understanding the biological variables contributing to psychiatric disorders has highlighted sex-based differences in the development and presentation of symptoms as well as in fundamental behavioral processes [1–4]. Understanding the factors contributing to sex-specific vulnerability to neuropsychiatric disease is critical to developing treatments that are safe and effective for both sexes.

Sex-based differences in reward seeking and avoidance offer an ideal model to explore bias and strategy in a behavioral task while providing insight into the neurobiological basis of information encoding [5–9]. For example, even though females will self-administer opiates at higher rates than males [2], when given a choice between opiates and a high-fat reward they choose the non-drug reinforcer over the drug alternative at a higher rate than

males [10], clearly highlighting that sex differences do not manifest themselves as universal behavioral principles, but rather are a complex interaction between sex and environment. Further, work in rats has shown while females are more motivated to self-administer drug and non-drug rewards [2, 7, 11, 12], they are also more sensitive to punishment [13]. Numerous other studies have varied the magnitude, value, and probability of rewards highlighting that females are more risk averse than males [13–17]. Capturing this complexity necessitates behavioral tasks that can probe the balance in the subjective value of rewarding versus aversive stimuli and their antecedent cues.

We established a task—Multidimensional Cue Outcome Action Task (MCOAT)—that allows for quantitative assessment of multidimensional behavioral functions relevant to human decision making in mice. By combining negative reinforcement, punishment, and positive reinforcement we can dissociate action from stimulus valence. Mice are first trained to respond to discriminative stimuli that predict either positive (nose-poke delivers sucrose) or negative (nose-poke removes shock) reinforcement. In subsequent trials both cues are presented concurrently and mice decide to respond to receive a reward or avoid a negative

¹Department of Pharmacology, Vanderbilt University School of Medicine, Nashville, TN 37232, USA; ²Vanderbilt Brain Institute, Vanderbilt University School of Medicine, Nashville, TN 37232, USA; ³Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA; ⁴Department of Psychiatry and Behavioral Sciences, Vanderbilt University School of Medicine, Nashville, TN 37232, USA and ⁵Vanderbilt Center for Addiction Research, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

Correspondence: Erin S. Calipari (erin.calipari@vanderbilt.edu)

These authors contributed equally: Munir Gunes Kutlu, Jennifer E. Zachry

Received: 3 September 2019 Revised: 23 April 2020 Accepted: 27 April 2020

Published online: 6 May 2020

outcome. In the last phase, aversive stimuli are delivered concurrently with rewards (punishment) and mice must inhibit responding. We apply this approach to demonstrate latent sex-specific behavioral strategies that are exposed at times of conflict or uncertainty. Next, we show how this task can be used to parse the behavioral processes that cellular populations in the brain modulate.

METHODS AND MATERIALS

See Supplementary methods for more detailed methodology.

Animals

Male and female 6- to 14-week-old C57BL/6J ($N = 75$; Jackson Laboratories; Bar Harbor, ME; #000664) or D1-Cre ($N = 13$; Jackson Laboratories; #030329) mice were housed five per cage. All animals were maintained on a reverse 12 h dark-12h light cycle and were food restricted to 90% of free-feeding weight. All experiments were conducted in accordance with the guidelines of the Institutional Animal Care and Use Committee at Vanderbilt University School of Medicine.

Multidimensional Cue Outcome Action Task (MCOAT)

Mice were trained/tested in Med Associates operant conditioning chambers (St. Albans, VT) (Fig. 1).

Phase 1: positive and negative reinforcement

Positive reinforcement: Mice were trained on a fixed-ratio 1 (FR1) schedule of reinforcement to nose-poke for sucrose (10 μ L volume, 1 mg sucrose; Fisher Scientific; Fig. 2a). Correct responses resulted in sucrose port illumination (5 s) and 1 s sucrose delivery. An auditory discriminative stimulus (S^{d1})—white noise or 2.5 kHz tone (counterbalanced)—was presented for the entirety of the session. Mice were moved to the next phase when they responded on the active NP > 80 times in a 1 h session.

Negative reinforcement: Mice were trained to nose-poke on the opposite, non-sucrose-paired nose-poke for negative reinforcement during 1 h sessions. Task order was counterbalanced. A second auditory discriminative stimulus (S^{d2})—tone or white noise, counterbalanced—was presented on a variable interval 30 s (VI30) schedule. The S^{d2} came on for 30 s after which a series of shocks (1.0 mA, 0.5 s) were delivered (15 s inter-stimulus interval, 20 shocks total). Correct responses during S^{d2} terminated S^{d2} and ended the trial, preventing shocks delivery. Correct responses made after shocks commenced terminated shocks. Unlike the positive reinforcement phase, discrete cues were used to signal the presence of an outcome to be removed. Acquisition criteria was defined as receiving fewer than 25% of total possible shocks in a session.

Phase 2a: limited discrimination and conflict

Limited discrimination pretraining. Animals underwent three sessions of discrimination training to ensure that they were using the antecedent cues (S^{d1} OR S^{d2}) to guide responses. S^{d1} and S^{d2} were presented in random order and equal proportion. Active/correct responses during S^{d1} initiated sucrose delivery and terminated S^{d1} . Response on the opposing nose-poke during S^{d2} terminated S^{d2} and ended the trial. Failure to make an active response during the 30 s duration of the S^{d2} resulted in a single shock.

Discrimination and conflict. Mice were trained in one session per day for three consecutive days. The test session consisted of both discrimination trials (80% of trials) and conflict trials (20% of trials) in the same session. Discrimination trials were identical to those described above. In conflict trials, mice were presented with a compound cue (S^{d1+2}) for 30 s. Depending on their response,

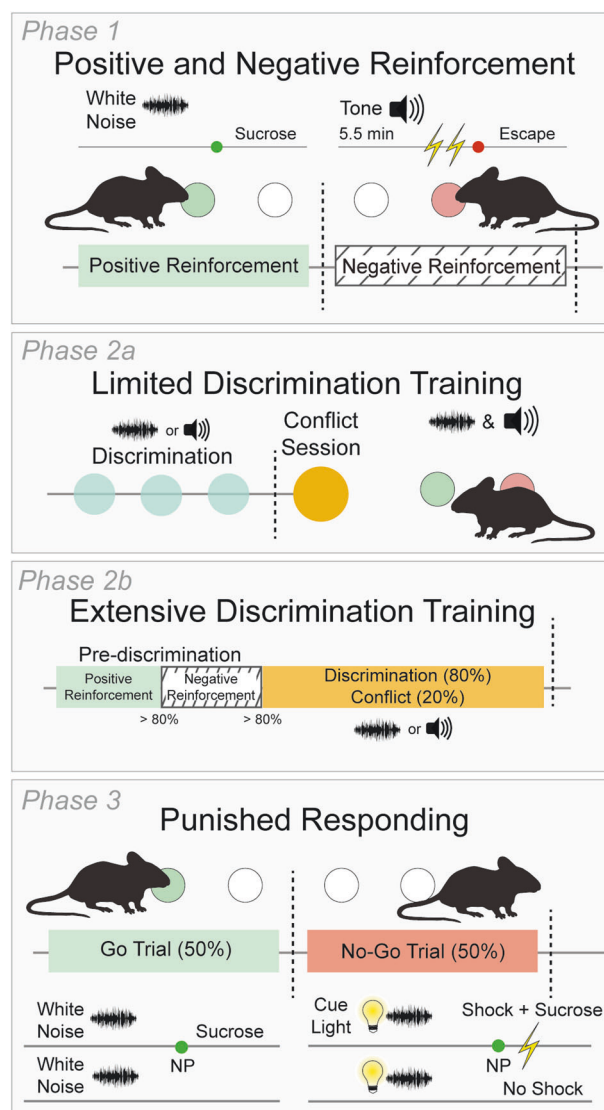


Fig. 1 Schematic of the Multidimensional Cue Outcome Action Task (MCOAT). In Phase 1, of the MCOAT mice are first trained in positive reinforcement on an FR1 schedule. A discriminative stimulus (S^{d1} , white noise or tone counterbalanced between animals and conditions) was presented for the duration of the session indicating that nose-pokes on a defined side (either left or right) were reinforced by sucrose delivery. In the second component of Phase 1, mice acquired negative reinforcement. In these trial-based sessions a separate auditory S^{d2} was presented to indicate that nose-poking (on the opposite nose-poke) was reinforced by the removal of a series of footshocks. Following acquisition mice transitioned to Phase 2a which had a discrimination phase (80% of trials), where each S^{d1} is randomly presented and animals are required to emit the correct operant response: poke for sucrose or poke for shock removal—depending on the S^{d1} presented. In the remaining trials (20%), S^{d1} and S^{d2} are presented simultaneously (S^{d1+2}) and animals have the option to nose-poke to obtain a sucrose reward or nose-poke to avoid shock. Thus, intrinsic response biases can be assessed. In Phase 2b, mice are trained until they meet a discrimination criterion of >70% and thus, have extensive experience with the discrimination/conflict portion of the task to understand how response bias changes with training experience. Finally, in Phase 3, animals are trained that a compound cue predicts punishment. Briefly, in 50% of the trials S^{d1} is presented and predicts positive reinforcement. In the remaining 50%, S^{d1} is presented with a secondary S^{d3} (S^{d1+3}) that predicts that a nose-poke will result in the delivery of sucrose and a footshock simultaneously. Together this behavioral task allows for the assessment across a wide range of approaches within individual animals.

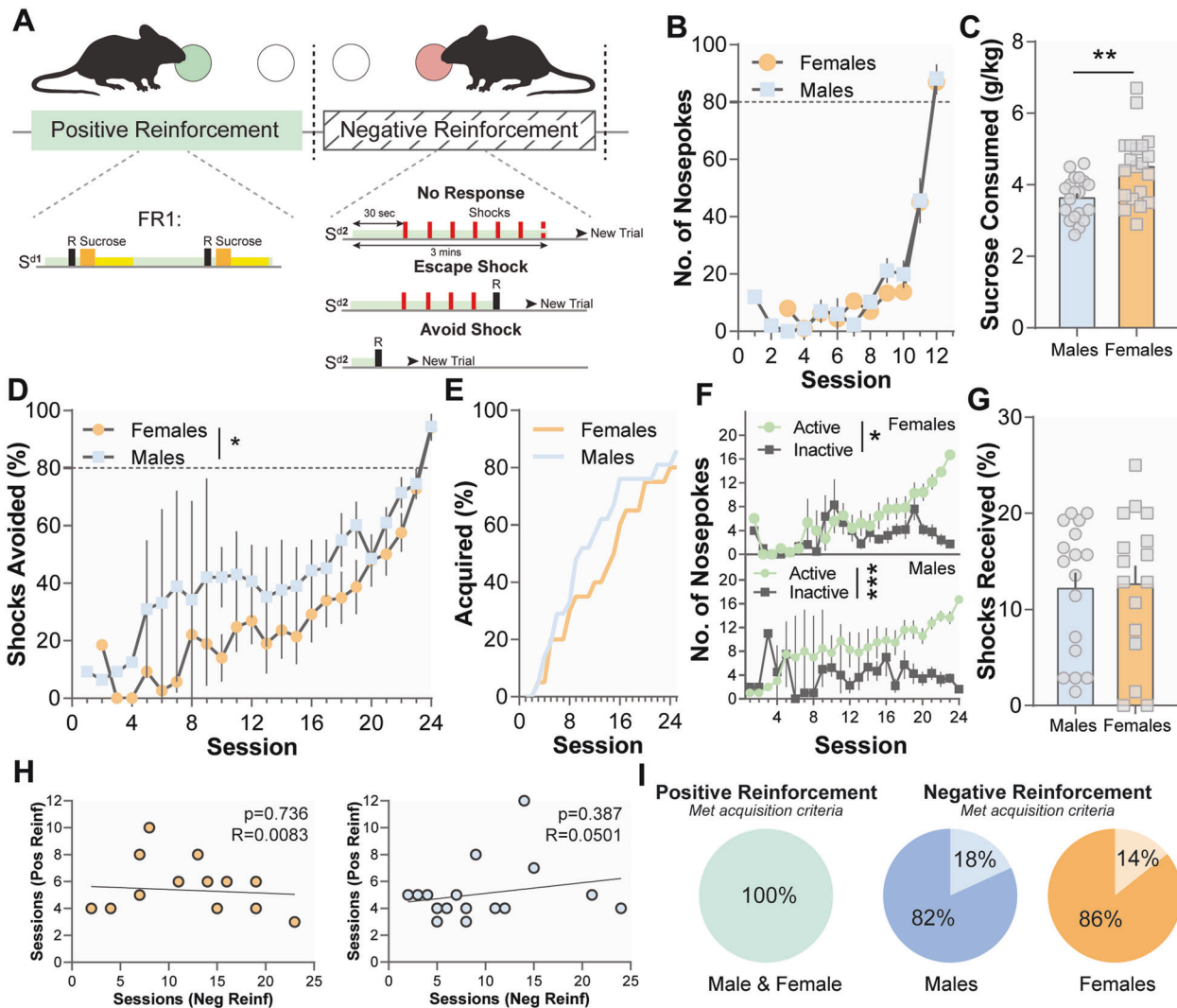


Fig. 2 Sex differences in reinforcement learning for positive and negative reinforcement. **a** Schematic of reinforcement schedules. Mice were trained on positive reinforcement and negative reinforcement. **b** Males ($n = 17$) and females ($n = 16$) learned positive reinforcement at similar rates. **c** Once acquisition criteria were met, females consume more sucrose than males. **d** Males acquired negative reinforcement at a faster rate than females. **e** Survival curve showing sex differences in the acquisition of negative reinforcement. **f** Both males and females showed differences in the number of active and inactive nose-pokes. Male and female mice that acquired the task did not differ in performance showing that the effect is selective to acquisition. **g** There was no difference in the number of shocks received once males and females reached criterion. **h** There is no relationship between the number of sessions to criterion for positive and negative reinforcement sessions in males or females showing that the task phases are independent measures. **i** Pie charts showing the percentage of animals that completed each component of phase I of the MCOAT. Data represented as mean \pm S.E.M. * $p < 0.05$.

mice received one of three possible outcomes: (1) failure to respond resulted in a single footshock, (2) responses on the sucrose port resulted in sucrose + footshock, and (3) responses on the negative reinforcement port resulted in shock avoidance.

Phase 2b: extensive discrimination and conflict. Mice underwent a 15-min pre-discrimination positive reinforcement session, a 15 min pre-discrimination negative reinforcement session (0.3 mA, 0.5 s shock), and a 1 h discrimination/conflict session. Mice received the shock pre-discrimination trials first to ensure they would still respond for sucrose before moving into the next phase. Mice that responded in $>80\%$ of the trials moved onto the next session. Mice were trained daily in discrimination until they reached a criterion of $>70\%$ correct.

Phase 3: punished responding
Positive reinforcement trials (50% of trials). Mice were presented S^{d1} and had 30 s to nose-poke on the active poke for sucrose

delivery. S^{d1} and the trial were terminated following an active response or at the end of the 30 s.

Punished trials (50% of trials). S^{d1} and S^{d3} (a house light) were presented concurrently. Active responses resulted in sucrose delivery and a single footshock. Shock intensity was increased over the course of 9 sessions (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.75, 1.0, and 1.5 mA).

Shock sensitivity
 Animals received randomly selected shocks of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.75, 1.0, or 1.5 mA with variable ITI of 30, 45, or 60 s. Vocalization (non-ultrasonic) and motor responses were scored. Vocalization was scored as 1 if the subject vocalized and 0 if the subject did not vocalize in the session. Motor responses were scored as 1 if the subject ran, 2 if the subject hopped (4 paws off the ground), 3 if the subject ran and hopped, and 0 if the subject did not move.

Chemogenetic inhibition experiments

Mice ($n = 13$; 6 males and 7 females) were positioned in a stereotaxic frame (Kopf Instruments) under isoflourane anesthesia. A 10-mL Nanofil Hamilton syringe (WPI) with a 34-gauge beveled needle was used to infuse AAV2/hSyn-DIO(Gi)-hM4Di-mCherry (Addgene #44362) into the NAc (bregma coordinates: anterior/posterior, +1.4 mm; medial/lateral, +1.5 mm; dorsal/ventral, -4.3 mm; 10° angle) of D1-Cre mice, thus achieving inhibitory DREADD expression in D1 MSNs in the NAc. Virus was infused at 50 nL/min for a total of 500 nL bilaterally. Animals recovered for four weeks before commencing behavioral experiments.

Mice were tested in Phase 1 and Phase 2a of the MCOAT. Clozapine N-oxide (CNO; 5 mg/kg) or saline was injected IP 30 min prior to behavioral testing to inhibit D1 MSNs during each discrete phase [18]. An experimental timeline denoting sessions where CNO was administered is presented in Fig. 6b. Mice were injected with CNO (or saline as a control) before the first and the second trial of positive and negative reinforcement to determine how this affected acquisition curves. Once mice had acquired, CNO/saline was administered during the last two sessions to determine the effects of D1 MSN inhibition on ongoing performance. During the discrimination phase, mice were given two drug-free discrimination sessions and received CNO or saline injections during the next two sessions. Finally, all mice received two conflict trials where they received the CNO/saline injections in a counterbalanced manner.

Histology

Subjects were deeply anaesthetized with an intraperitoneal injection of Ketamine/Xylazine (100 mg/kg/10 mg/kg) and transcardially perfused with 10 mL of PBS solution followed by 10 mL of cold 4% PFA in 1× PBS. Brains were sectioned at 35 μm on a freezing sliding microtome (Leica SM2010R) and fluorescent images were taken using a Keyence BZ-X700 inverted fluorescence microscope (Keyence; Fig. 6a). One mouse was removed from analysis because it only showed unilateral expression of the inhibitory DREADDs.

Analysis. For positive and negative reinforcement, the total sucrose and shock responses were analyzed using unpaired *t* tests. Mann–Whitney *U* test was used when the number of sessions to criterion was not equal between subjects. Discrimination and conflict responses were analyzed using two-way ANOVA (Trial Type × Sex). We employed a mixed Repeated Measures ANOVA for the Punished Responding and Shock Sensitivity experiments. For the DREADDs experiments, we calculated the Trial × Drug interactions using a mixed Repeated Measures ANOVA (for positive and negative reinforcement) and a two-way ANOVA (for discrimination and conflict). Power calculations were done to ensure sufficient power and adequate group sizes (Supplementary Table 1). We also determined the parameters of response bias (Log *b*) and discrimination (Log *d*), as described previously [19, 20]. Both terms use a logarithmic scale for the multiplication of the ratio between correct and incorrect responses during two different trial types:

Log *d*. Log *d* is the ratio between the number of correct and incorrect Sucrose and Shock trials, which results in a negative (no discrimination) or a positive (successful discrimination):

$$\text{Log } d = 0.5 \times \log\left[\frac{((\text{Sucrose}_{\text{correct}} + 0.5) \times (\text{Shock}_{\text{correct}} + 0.5))}{((\text{Sucrose}_{\text{incorrect}} + 0.5) \times (\text{Shock}_{\text{incorrect}} + 0.5))}\right]$$

Log *b*. Log *b* is calculated as the ratio between the number of correct Sucrose and incorrect Shock versus incorrect Sucrose and correct Shock trials, which results in either a negative (bias toward avoidance) or a positive value (bias toward sucrose) or a 0

(no bias):

$$\text{Log } b = 0.5 \times \log\left[\frac{((\text{Sucrose}_{\text{correct}} + 0.5) \times (\text{Shock}_{\text{correct}} + 0.5))}{((\text{Sucrose}_{\text{incorrect}} + 0.5) \times (\text{Shock}_{\text{correct}} + 0.5))}\right]$$

RESULTS

Phase 1: females show increased measures of positive reinforcement and a decreased learning rate for negative reinforcement

Although females consumed more sucrose (Fig. 2c; $t(38) = 2.603$, $p = 0.0019$), there was no difference in the positive reinforcement learning rate in males and females (Fig. 2b; $p > 0.05$). In contrast, males avoided a larger percentage of shocks during the acquisition phase as compared with females (Fig. 2d; Mann–Whitney $U = 164$, $p = 0.0164$). Both males (Mann–Whitney $U = 96.50$, $p < 0.0001$) and females (Mann–Whitney $U = 150.5$, $p = 0.0114$) showed differences in the number of active vs inactive nose-pokes (Fig. 2f). The percentage of male and female mice that completed the task did not differ (Fig. 2e, i; $X^2(22) = 14.97$, $p > 0.05$) and there was no difference in the number of shocks male and female mice received once they reached criterion (Fig. 2g; $t(30)$, 0.5210, $p > 0.05$). These differences were not driven by differences in body weight (males weighed more than females; Supplementary Fig 1) or differences in consumption at baseline between the sexes (Supplementary Fig. 2).

One of the major components of this task is to dissociate different behavioral strategies, which depends on each phase of the task being independent from one another. Indeed, we found no correlation between number of sessions to criterion for positive and negative reinforcement in males or females (Fig. 2h; $p > 0.05$).

Phase 2a: females' responses are biased toward avoiding negative outcomes when conflicting information is presented

Mice received a limited number of discrimination training sessions to confirm they were using the **S^d**, followed by a test session where 80% of the trials were discrimination trials and 20% were conflict trials (Fig. 3a). Both sexes showed similar levels of discrimination (Fig. 3b; Sex × Trial Type Interaction, $F(1,14) = 2.885$, $p > 0.05$; Sex main effect, $F(1,14) = 0.028$, $p > 0.05$) and Log *d* (Fig. 3c; $t(14)$, 0.5797, $p > 0.05$). There was no main effect of Trial Type ($F(1,14) = 0.526$, $p > 0.05$), indicating that animals completed similar numbers of sucrose/avoidance trials.

During conflict trials, there was an interaction between Sex and Trial Type (Fig. 3b; $F(1,24) = 7.410$, $p = 0.0119$) and a sex difference for Log *b* (Fig. 3c; $t(14) = 2.159$, $p = 0.0487$), demonstrating that male and female mice show differential biases. While female mice chose to avoid shocks over sucrose, male mice did not show a bias. In line with this, the difference between the number of sucrose and avoidance responses was significant for females (Fig. 3b; $t(6) = 4.816$, $p = 0.0030$) but not males ($t(6) = 0.2448$, $p > 0.05$). We also found no correlation between Log *b* and Log *d* values with sessions to complete the positive or negative reinforcement task (Fig. 3d; $p > 0.05$). Overall, female mice—but not males—have an intrinsic bias toward avoiding aversive outcomes.

Phase 2b: females response bias does not change over extensive training

A second group of mice underwent extensive discrimination training (Fig. 4a). The goal was to examine: (1) the progression of discrimination learning and behavioral bias and (2) behavioral bias during conflict when animals reach a set level of discrimination and are familiar with the task (>70% correct for both sucrose and shock trials). We found an interaction for Discrimination/Bias and Trial for males (Fig. 4b; $F(40, 256) = 1.718$, $p < 0.01$) but not for females (Fig. 4b; $F(32, 198) = 1.076$, $p > 0.05$). Although we could not test the interaction between Sex and Trials due to faster discrimination learning in males resulting in unequal number of

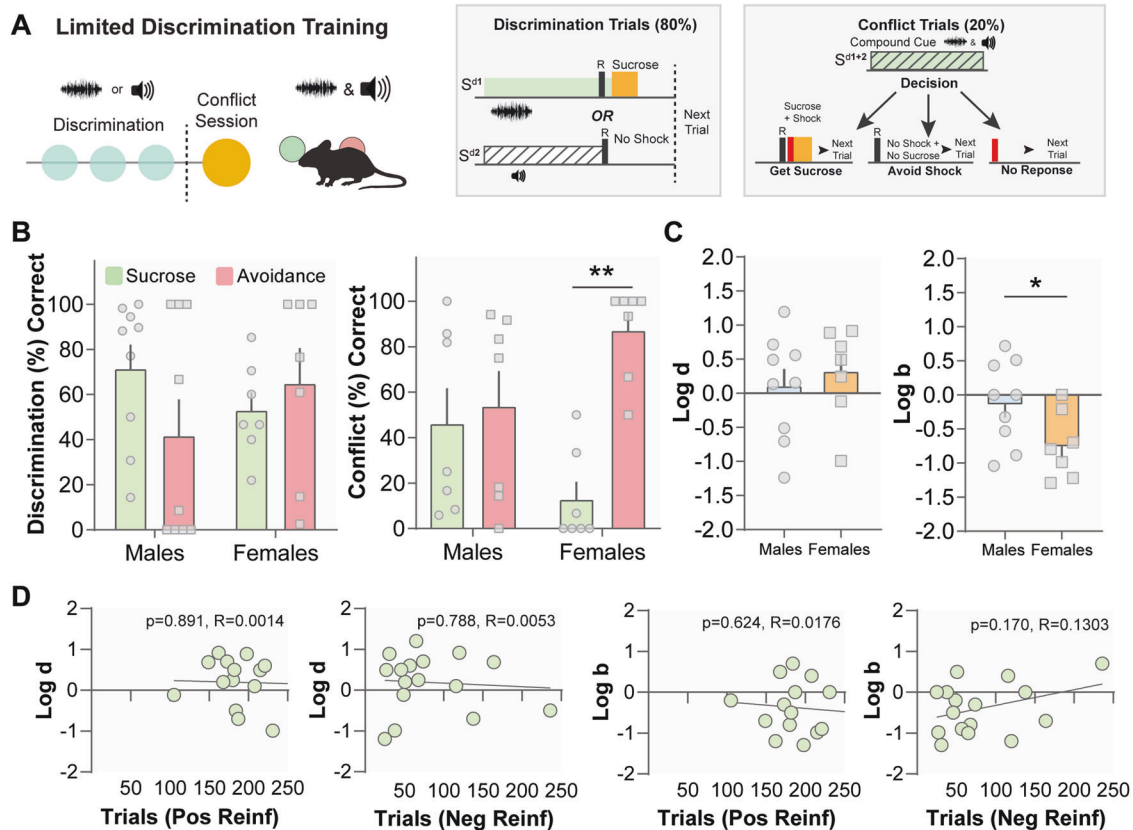


Fig. 3 Females are biased toward shock avoidance. **a** In limited discrimination, mice are given three discrimination sessions. In the fourth session, 80% of the trials remain discrimination trials and 20% of the trials are conflict trials, in which both S^ds are presented simultaneously. **b** Males ($n = 9$) and females ($n = 7$) show comparable levels of discrimination on the final discrimination session. **b** In conflict trials, females responded for shock avoidance more than sucrose. For the males, these numbers did not differ. **c** There were no sex differences in Log d , showing both males and females discriminated the cues similarly. However, females showed bias toward shock avoidance, while males do not demonstrate bias. **d** There is no relationship between Log b and Log d values with the number of sessions to meet the criteria for positive or negative reinforcement. Data represented as mean \pm S.E.M. * $p < 0.05$, ** $p < 0.01$.

data points between groups, there was a difference in Log d (Mann–Whitney $U = 374$, $p < 0.001$) and Log b (Mann–Whitney $U = 165$, $p < 0.0001$) values between males and females. There was no difference in sessions to criterion between sexes (Fig. 4c; $t(17) = 0.6476$, $p > 0.05$). Furthermore, once male and female mice reached discrimination criterion they showed comparable levels of discrimination (Fig. 4d; Supplementary Figs. 3 and 4; Sex \times Trial Type Interaction, $F(1,17) = 0.8253$, $p > 0.05$). At the end of discrimination training both sexes showed a response bias toward choosing avoidance over sucrose during the conflict trials (Fig. 4d; Sex \times Trial Type Interaction, $F(1,34) = 1.387$, $p > 0.05$; Trial Type Main Effect, $F(1,34) = 50.33$, $p < 0.0001$, $p < 0.0001$; $t(9) = 3.807$, $p < 0.01$; $t(8) = 3.391$, $p < 0.01$), but there was no interaction between Sex and Trial for Log d ($F(1, 17) = 0.4983$, $p > 0.05$). The main effect of Trial was significant (Fig. 4e; $F(1, 17) = 39.97$, $p < 0.0001$). Similarly, sex differences in Log d ($t(17) = 0.3799$, $p > 0.05$) and Log b ($t(17) = 0.5572$, $p > 0.05$) disappeared following increased familiarity with the task (Fig. 4f). Finally, Log b and Log d values were negatively correlated for the limited discrimination phase ($p < 0.01$, $R = 0.4794$), indicating that animals who showed better discrimination were more inclined to avoid footshocks over obtaining sucrose (Supplementary Fig. 5).

Phase 3: female mice are more sensitive to punishment
Punishers function to decrease rates of responding [21] and there was a difference between males and females in the number of sucrose responses during punishment (Fig. 5c; Sex \times Shock Intensity Interaction, $F(8, 72) = 0.3219$, $p = 0.9552$; Sex Main Effect,

$F(1, 9) = 5.157$, $p = 0.0493$; Shock Intensity Main Effect, $F(2.930, 26.37) = 15.16$, $p < 0.0001$) as well as for the number of sucrose+shock responses (Fig. 5c; Sex \times Shock Intensity Interaction, $F(8, 72) = 1.259$, $p > 0.05$; Sex Main Effect, $F(1, 9) = 4.644$, $p = 0.0596$; Shock Intensity Main Effect, $F(2.074, 18.67) = 41.02$, $p < 0.0001$) between varying shock intensities. In addition, the Sex \times Trial Type interaction is significant (Fig. 5h; $F(1, 81) = 16.03$, $p < 0.001$) as well as the main effect of Trial Type for males ($F(1, 10) = 10.90$, $p = 0.0080$) but not for females ($F(1, 8) = 3.812$, $p = 0.0867$) suggesting both groups learned to differentiate between sucrose and punished trials; however, females were more sensitive to the effects of punishers. We also computed response inhibition 50 curves (RI₅₀) for each animal to determine the shock intensity value which caused a 50% reduction in behavioral responding (Fig. 5d, f). Females required lower shock intensities to reduce responses in non-shock (Fig. 5e; Mann–Whitney $U = 5$, $p = 0.0480$) but not fewer shock trials (Fig. 5g; $t(9) = 1.971$, $p > 0.05$). Sucrose to shock RI₅₀ ratios did not differ between sexes (Fig. 5i; $t(9) = 0.9758$, $p > 0.05$). Overall, females are more sensitive than males to aversive outcomes regardless of whether the response is active or requires response inhibition (Fig. 5a).

Females are less sensitive to footshock

There was a sex difference in motor response (Fig. 5b; Sex \times Shock Intensity Interaction, $F(4,35) = 2.802$, $p = 0.0406$; Sex Main Effect, $F(1,35) = 1.830$, $p > 0.05$; Shock Intensity Main Effect, $F(4,35) = 48.33$, $p < 0.0001$) and in vocalization (Fig. 5b; Sex \times Shock Intensity Interaction, $F(4,35) = 6.991$, $p = 0.0003$; Sex Main Effect,

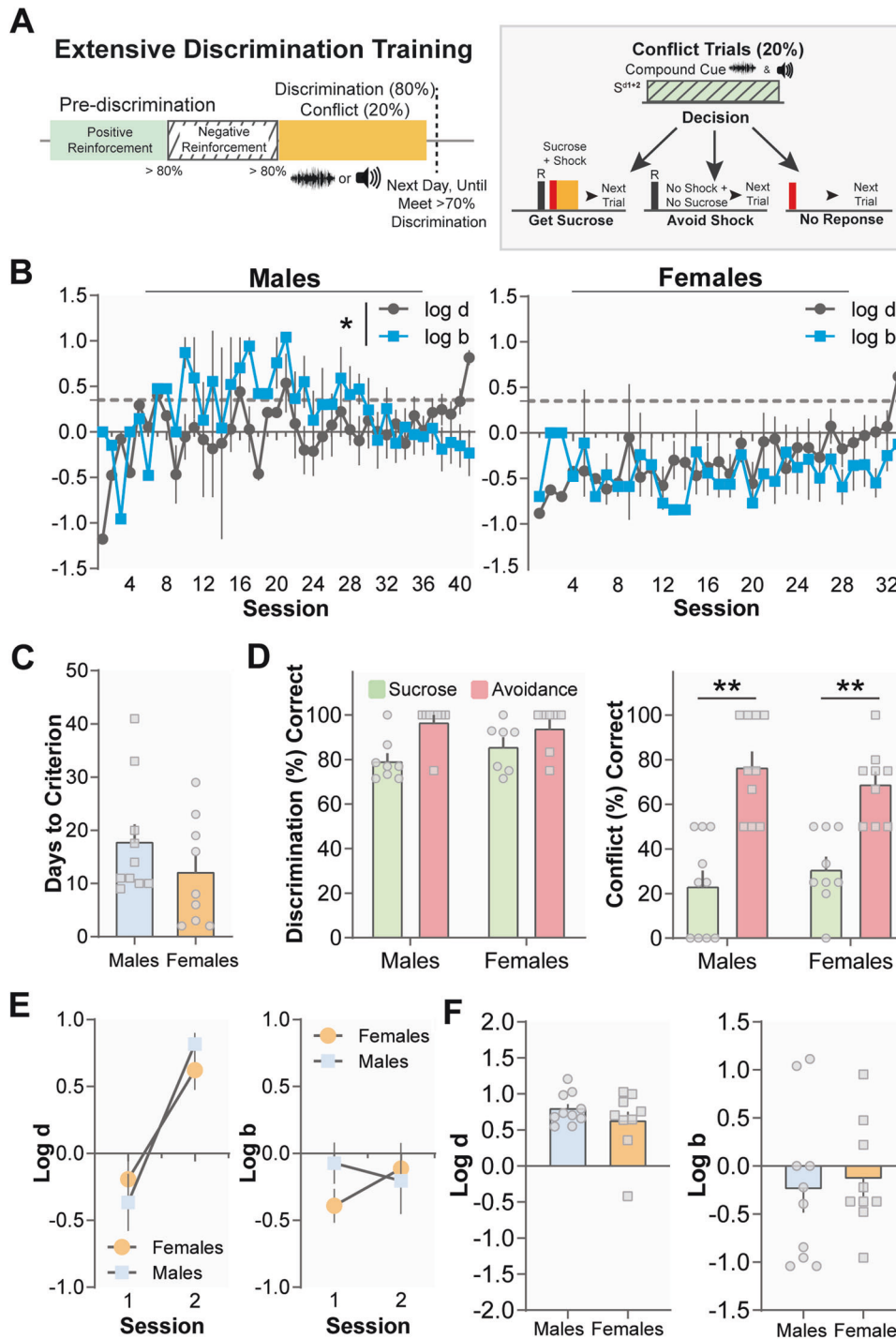


Fig. 4 Extensive training on the MCOAT does not alter female bias toward avoiding aversive stimuli. **a** Schematic of extensive discrimination and conflict task. **b** Males, but not females, significantly changed their bias over extensive experience with the task. **c** There was no difference in the number of days to criterion between sexes. **d** Once both groups met discrimination criterion, they demonstrated similar levels of discrimination. Following extended discrimination training, males and females both showed a response bias for shock avoidance in conflict trials. **e** Both males and females showed increased *Log d* values with extended discrimination. **f** At the end of the extended discrimination phase, males and females show similar *Log d* and *Log b* values. Data represented as mean \pm S.E.M. ****** $p < 0.01$.

$F(1,35) = 27.58, p < 0.0001$; Shock Intensity Main Effect, $F(4,35) = 111.3, p < 0.0001$) to different shock intensities. However, Bonferroni-corrected *t*-tests showed that only males showed higher motor response to 0.30 mA shock ($p = 0.010$). The direction of this effect suggests male mice are more sensitive to lower shock intensities, while females are more sensitive to the effects of shock on reinforcement/punishment.

D1 MSNs modulate stimulus-specific learning

An advantage of the MCOAT is that animals learn both positive and negative reinforcement. Because the action is the same (i.e., reinforcement), if a cellular population controls reinforcement the effects should be similar between the two task types. If the population controls stimulus processing (i.e., stimulus valence) differences between the two task types will emerge.

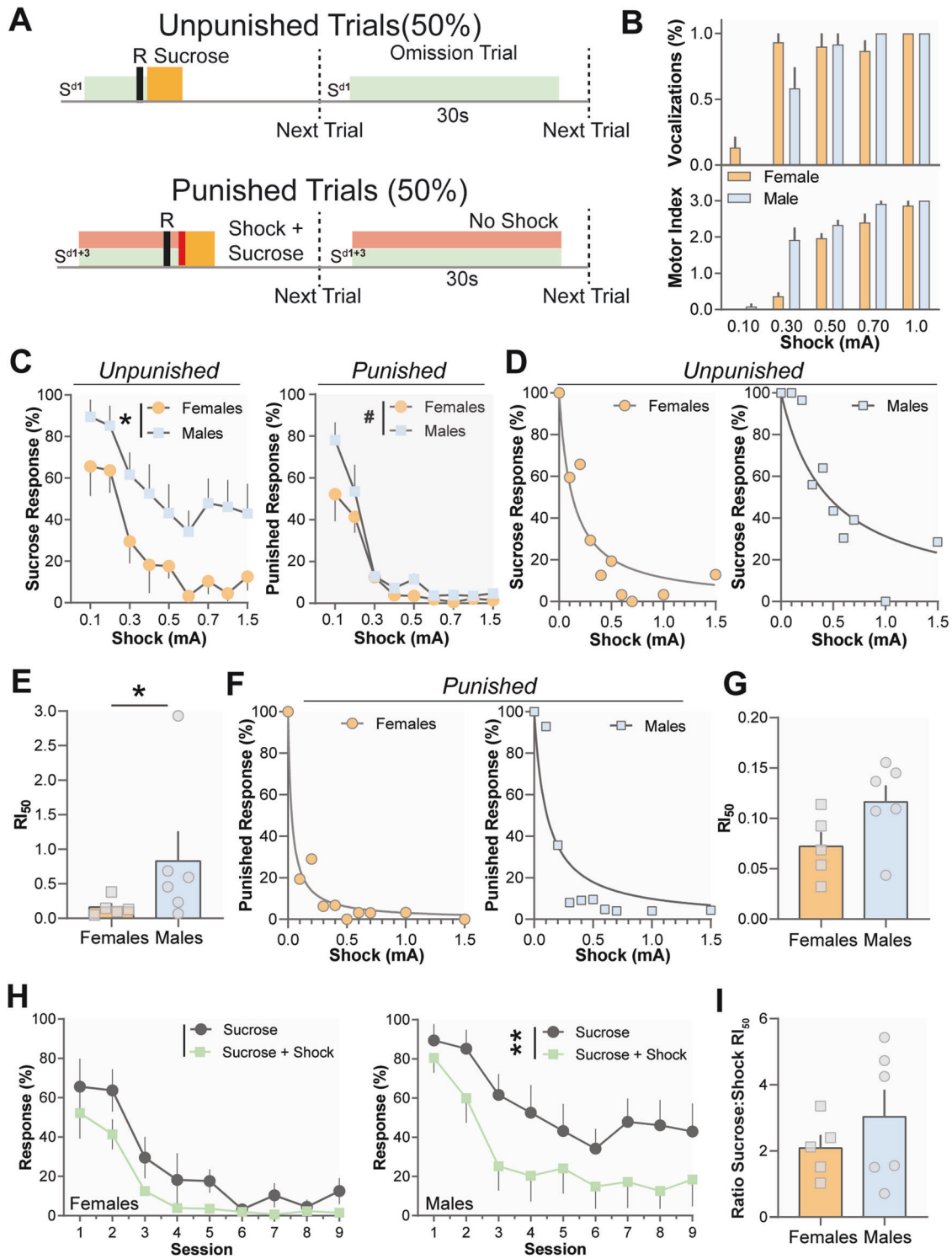


Fig. 5 Females are more sensitive to punishment. **a** Schematic of punished responding. **b** Males ($n = 4$) were more sensitive than females ($n = 5$) to unsignaled shocks of varying intensities. Females were less sensitive to shock presentation as measured by motor response and vocalizations. **c** Males ($n = 7$) respond more during both unpunished (left) and punished (right) trials as compared with females ($n = 5$) over increasing shock intensities. **d** Representative response inhibition 50 curves (RI_{50}) for sucrose responding for males and females. **e** Males show a higher RI_{50} than females for sucrose responding indicating that more shock intensity was necessary to reduce their response rates. **f** Representative response inhibition 50 curves (RI_{50}) for punished responding for males and females. **g** There was a trend toward higher RI_{50} in males during punishment. **h** When analyzed separately males showed a decreased response rate during Sucrose + Shock trials while females did not. **i** The relationship between unpunished and punished responding (sucrose responding RI_{50} /punished responding RI_{50}) was not different between sexes. Data represented as mean \pm S.E.M. * $p < 0.05$, ** $p < 0.01$, #0.0596.

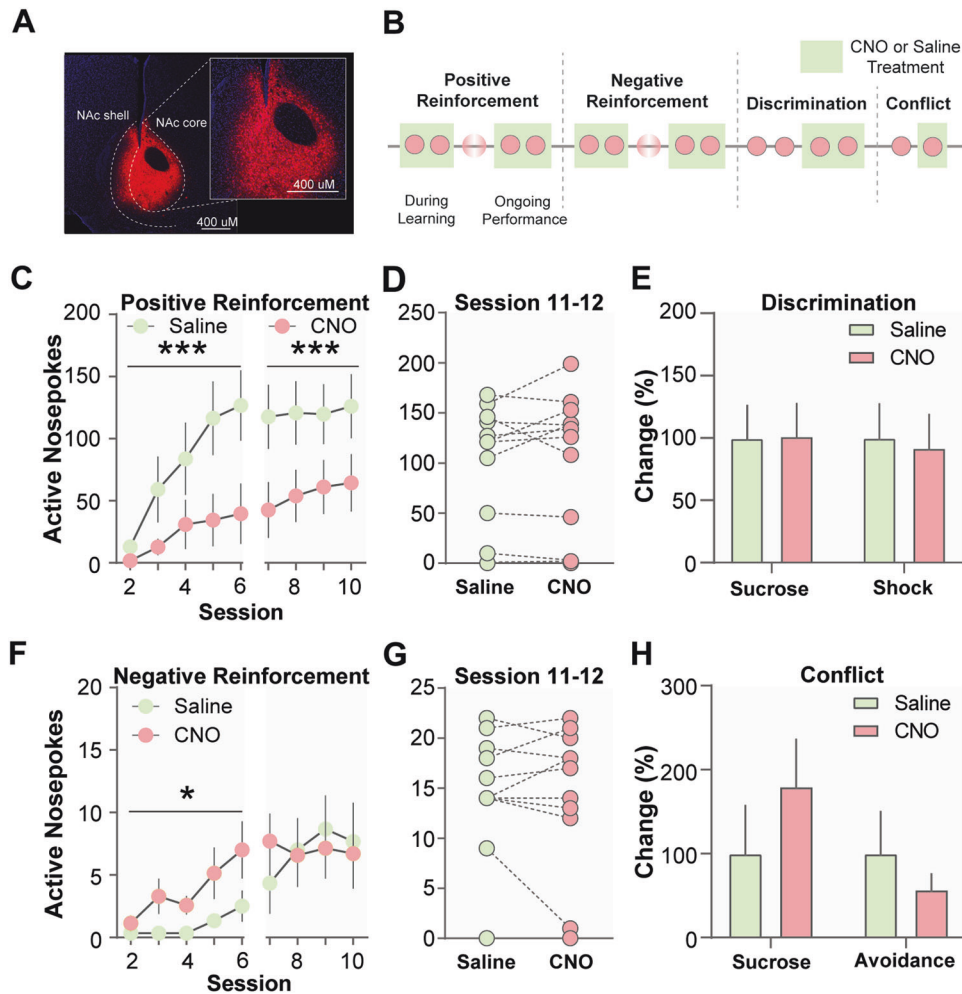


Fig. 6 Chemogenetic inhibition of D1 MSNs in the NAc disrupts positive reinforcement and enhances negative reinforcement learning. **a** Representative image of viral expression of DREADDs (hm4Di—inhibitory) in the NAc core of D1-Cre mice. **b** Schematic of experimental design. **c** Chemogenetic inhibition of NAc D1 MSNs during positive reinforcement learning reduced active responses. **d** After the learning criterion was met, inhibition of D1 MSNs did not affect task performance. **f** D1 MSN inhibition enhanced acquisition of negative reinforcement. **g** D1 MSN inhibition did not affect post-training performance. **e** D1 MSN inhibition did not alter discrimination between sucrose and avoidance responses. **h** D1 MSN inhibition did not change response bias during conflict trials. Data represented as mean \pm S.E.M. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.0001$.

Chemogenetic inhibition of NAc D1 MSNs decreased the number of correct responses during positive reinforcement learning (Fig. 6c; Drug \times Session Interaction, $F(8,99) = 0.5122$, $p > 0.05$; Drug Main Effect, $F(1,99) = 32.40$, $p < 0.0001$; Session Main Effect, $F(8,99) = 3.422$, $p = 0.0016$). However, there was no effect after animals learned the task (Fig. 6d; $t(12) = 0.8037$, $p > 0.05$). For negative reinforcement there was also a main effect of CNO during acquisition of negative reinforcement ($F(1,11) = 5.119$, $p = 0.0449$); however, learning was enhanced rather than inhibited. Similar to experiments above, the effects were specific to learning and CNO injections did not affect performance once the task was learned (Fig. 6g; $t(12) = 0.3828$, $p > 0.05$).

Giving further support to the hypothesis that D1 MSNs play a critical role in acquisition, but not ongoing performance was data showing that CNO injections had no effect on discrimination (Fig. 6e; Drug \times Trial Type Interaction, $F(1,12) = 0.7692$, $p > 0.05$; Drug Main Effect, $F(1,12) = 0.7692$, $p > 0.05$; Trial Type Main Effect, $F(1,12) = 1.831$, $p > 0.05$), or on response bias during conflict trials (Fig. 6h; Drug \times Trial Type Interaction, $F(1,12) = 2.532$, $p > 0.05$; Drug Main Effect, $F(1,12) = 0.2453$, $p > 0.05$; Trial Type Main Effect, $F(1,12) = 1.708$, $p > 0.05$). Together, these data show the effects of D1 MSNs are stimulus-specific and are selective to learning.

DISCUSSION

Emerging evidence—including work presented here—has highlighted that biological sex itself is not a behavioral determinant, but rather a complex variable interacting with environmental factors and experience to drive behavior, thus requiring studies that allow for an understanding of the behavioral factors that underlie sex-specific strategies [22–25]. Here, we present a complex behavioral task (MCOAT) that further highlights the complexities of sexually dimorphic behaviors (Supplementary Fig. 6). First, we showed that female mice self-administer higher levels of sucrose but acquire negative reinforcement at a slower rate; however, in situations where positive and negative stimuli are presented together (conflict or punishment), females favor avoiding aversive outcomes over seeking rewards. Together, fundamental differences in basic behavioral strategies between the sexes—specifically in regard to the balance of positive and negative outcomes—point to critical factors that may guide behavior in females and may underlie the differences in development and progression of psychiatric disease states [25].

Decision making is a process in which various external and internal processes are in play to ensure homeostasis between positive and negative outcomes [26]. Here, we present evidence

demonstrating that female mice have an intrinsic bias toward avoiding negative outcomes over obtaining rewards, while male mice showed a similar bias only when they have extensive experience with cue-outcome contingencies suggesting male mice seek out rewards without considering aversive outcomes to the same extent. Importantly, the behavioral bias that female mice showed was not due to a differential ability to perform the task, as at the end of the negative reinforcement and discrimination phases, both sexes exhibited similar levels of performance. It is interesting to note that the response in males changed with extended training, suggesting that learning may play a role in this effect and that females apply strategies resulting in response bias toward shock avoidance sooner than the males. Indeed, there are several studies showing that females adopt risk-averse strategies sooner than males [14, 15, 27].

Further, we show that sex differences in learning are dependent on stimulus valence whereby females learn behavior reinforced by a negative stimulus (shock) at a slower rate than males, while learning rates are the same when reinforced by a positive stimulus (sucrose). Importantly, the motoric action associated with both of the stimuli are identical, allowing us to rule out differences in movement as a driver of this difference. Similarly, correct responses on both trial types results in a positive outcome (delivery of sucrose, or removal of shock), which allows for specific determination that these effects are driven by stimulus valence rather than outcome valence. Indeed, the ability to look in the same experimental subject over tasks with divergent stimuli (sucrose vs. footshock) but convergent (nose-poke response) outcomes OR convergent stimuli (shock) with divergent behavioral outcomes (i.e. response vs response inhibition) is a major advantage of the MCOAT. In the initial phase of the task, we show that females self-administer more sucrose than males, which is in line with previous studies reporting increased reward seeking in females for both natural [28, 29] and drug reinforcers [7, 11, 30, 31]. However, these previous findings have led to the conclusion that females are more driven by positive outcomes. This indicates that simple behavioral tasks developed to test a single dimension of behavior may lead to incomplete conclusions about sex differences.

The MCOAT is not the first attempt to examine how outcomes are weighed in a reinforcement learning context. Indeed, several models of risky decision making have proven efficient in examining these behaviors in rats [13, 16, 32]. The MCOAT is designed to ask different questions as it utilizes negative reinforcement and punishment, which allows for the dissociation between differential effects of the behavioral action or stimulus processing. For example, Orsini et al. [13] alters the value of two potential outcomes—therefore every response is ultimately reinforced by a reward, either in the presence or absence punishment. The MCOAT is different in that it assesses responding in the face of a punisher as well as negative reinforcement—thus dissociating differences in behavioral action from stimulus processing itself. Finally, this model is in mice whereas virtually all motivated action and decision-making models are employed in rats, which has limited the work outlining genetically-defined cell populations and how they control discrete aspects of behavior.

To this end, we show how the MCOAT can be utilized to define the involvement of specific neural populations in behavioral control, by combining the task with chemogenetic approaches to inhibit D1 MSNs in the NAc. D1 MSNs in the NAc are thought to drive aspects of reward learning [18, 33, 34], however, it is not clear if this is due to their actions on motivated behavior (i.e., reinforcement/seeking) or stimulus processing. Our results showed that inhibition of D1 MSNs reduces positive and enhances negative reinforcement learning without affecting post-training performance. Thus, these results demonstrate that NAc D1 MSNs are involved in stimulus processing during learning, rather than reinforcement behavior itself.

Together, these experiments demonstrate that when combined with neural intervention techniques the MCOAT allows for clear dissociation of neural encoding of stimulus processing, actions, and response discrimination and bias to isolate these critical aspects of behavioral control. This behavioral procedure will be particularly powerful for examining the effects of many other external (e.g., stress, depression, and addiction) or internal (hunger and thirst) conditions on learned behavior and behavioral biases. Understanding these processes precisely is critically important to improving treatments for these conditions, especially in women where treatment efficacy is reduced and off-target and adverse consequences from medications are particularly high [35–37].

FUNDING AND DISCLOSURE

The authors declare no competing interests.

ACKNOWLEDGEMENTS

This work was supported by NIH grants DA042111, DA048931 to ESC, DA045103 to CAS, GM07628 to JEZ, MH065215 and DA047777 to ARJ, MH064913 and DA050410 to KCT, and DA048436 to AJL as well as by funds from the VUMC Faculty Research Scholar Award to MGK, the Vanderbilt Academic Pathways Fellowship to LJB, Brain and Behavior Research Foundation to MGK, ESC, and CAS, the Whitehall Foundation to ESC, and the Edward Mallinckrodt Jr. Foundation to ESC.

ADDITIONAL INFORMATION

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41386-020-0692-1>).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Altemus M, Sarvaiya N, Neill Epperson C. Sex differences in anxiety and depression clinical perspectives. *Front Neuroendocrinol.* 2014;35:320–30.
2. Becker JB, Hu M. Sex differences in drug abuse. *Front Neuroendocrinol.* 2008;29:36–47.
3. Calipari ES, Juarez B, Morel C, Walker DM, Cahill ME, Ribeiro E, et al. Dopaminergic dynamics underlying sex-specific cocaine reward. *Nat Commun.* 2017. <https://doi.org/10.1038/ncomms13877>.
4. Chowdhury, T. G., Wallin-Miller, K. G., Rear, A. A., Park, J., Diaz, V., Simon, N. W., & Moghaddam, B. Sex differences in reward-and punishment-guided actions. *Cognitive, Affective, & Behavioral Neuroscience*, 2019;19:1404–17.
5. Johnson AR, Thibeault KC, Lopez AJ, Peck EG, Sands LP, Sanders CM, et al. Cues play a critical role in estrous cycle-dependent enhancement of cocaine reinforcement. *Neuropsychopharmacology.* 2019. <https://doi.org/10.1038/s41386-019-0320-0>.
6. Becker JB, Chartoff E. Sex differences in neural mechanisms mediating reward and addiction. *Neuropsychopharmacology.* 2019;44:166–83.
7. Hu M, Crombag HS, Robinson TE, Becker JB. Biological Basis of Sex Differences in the Propensity to Self-administer Cocaine. *Neuropsychopharmacology.* 2004. <https://doi.org/10.1038/sj.npp.1300301>.
8. Dalla C, Shors TJ. Sex differences in learning processes of classical and operant conditioning. *Physiol Behav.* 2009. <https://doi.org/10.1016/j.physbeh.2009.02.035>.
9. Chowdhury, T. G., Wallin-Miller, K. G., Rear, A. A., Park, J., Diaz, V., Simon, N. W., & Moghaddam, B. Sex differences in reward-and punishment-guided actions. *Cognitive, Affective, & Behavioral Neuroscience*, 2019;19:1404–17.
10. Townsend EA, Negus SS, Caine SB, Thomsen M, Banks ML Sex differences in opioid reinforcement under a fentanyl vs. food choice procedure in rats. *Neuropsychopharmacology.* 2019. <https://doi.org/10.1038/s41386-019-0356-1>.
11. Lynch WJ, Carroll ME Sex differences in the acquisition of intravenously self-administered cocaine and heroin in rats. *Psychopharmacology.* 1999. <https://doi.org/10.1007/s002130050979>.
12. Lynch WJ, Roth ME, Carroll ME. Biological basis of sex differences in drug abuse: Preclinical and clinical studies. *Psychopharmacology.* 2002;164:121–37.
13. Orsini CA, Willis ML, Gilbert RJ, Bizon JL, Setlow B. Sex differences in a rat model of risky decision making. *Behav Neurosci.* 2016. <https://doi.org/10.1037/bne0000111>.

14. Ishii H, Onodera M, Ohara S, Tsutsui KI, Iijima T. Sex differences in risk preference and c-Fos expression in paraventricular thalamic nucleus of rats during gambling task. *Front Behav Neurosci*. 2018. <https://doi.org/10.3389/fnbeh.2018.00068>.
15. Van den Bos R, Jolles J, Van der Knaap L, Baars A, De Visser L. Male and female Wistar rats differ in decision-making performance in a rodent version of the Iowa Gambling Task. *Behav Brain Res*. 2012. <https://doi.org/10.1016/j.bbr.2012.07.015>.
16. Orsini CA, Setlow B. Sex differences in animal models of decision making. *J Neurosci Res*. 2017.
17. Jolles JW, Boogert NJ, van den Bos R. Sex differences in risk-taking and associative learning in rats. *R Soc Open Sci*. 2015. <https://doi.org/10.1098/rsos.150485>.
18. Calipari ES, Bagot RC, Purushothaman I, Davidson TJ, Yorgason JT, Peña CJ, et al. In vivo imaging identifies temporal signature of D1 and D2 medium spiny neurons in cocaine reward. *Proc Natl Acad Sci USA*. 2016. <https://doi.org/10.1073/pnas.1521238113>.
19. Branch MN. On the role of "memory" in the analysis of behavior. *J Exp Anal Behav*. 1977. <https://doi.org/10.1901/jeab.1977.28-171>.
20. Kangas BD, Berry MS, Branch MN. On the Development and Mechanics of Delayed Matching-to-Sample Performance. *J Exp Anal Behav*. 2011. <https://doi.org/10.1901/jeab.2011.95-221>.
21. Azrin NH. Punishment and recovery during fixed-ratio performance. *J Exp Anal Behav*. 1959;2:301.
22. Cahill L. Why sex matters for neuroscience. *Nat Rev Neurosci*. 2006;7:477–84.
23. Kiraly DD, Walker DM, Calipari ES, Labonte B, Issler O, Pena CJ, et al. Alterations of the host microbiome affect behavioral responses to cocaine. *Sci Rep*. 2016;6.
24. Siciliano CA. Capturing the complexity of sex differences requires multi-dimensional behavioral models. *Neuropsychopharmacology*. 2019;44:1997–8.
25. Zachry JE, Johnson AR, Calipari ES. Sex differences in value-based decision making underlie substance use disorders in females. *Alcohol Alcohol*. 2019;54:339–41.
26. Gu X, FitzGerald THB. Interoceptive inference: homeostasis and decision-making. *Trends Cogn Sci*. 2014. <https://doi.org/10.1016/j.tics.2014.02.001>.
27. Georgiou P, Zanos P, Bhat S, Tracy JK, Merchenthaler IJ, McCarthy MM, et al. Dopamine and stress system modulation of sex differences in decision making. *Neuropsychopharmacology*. 2018. <https://doi.org/10.1038/npp.2017.161>.
28. Brown KJ, Grunberg NE. Effects of environmental conditions on food consumption in female and male rats. *Physiol Behav*. 1996. [https://doi.org/10.1016/0031-9384\(96\)00020-0](https://doi.org/10.1016/0031-9384(96)00020-0).
29. Hong S, Flashner B, Chiu M, ver Hoeve E, Luz S, Bhatnagar S. Social isolation in adolescence alters behaviors in the forced swim and sucrose preference tests in female but not in male rats. *Physiol Behav*. 2012. <https://doi.org/10.1016/j.physbeh.2011.08.036>.
30. Middaugh LD, Kelley BM, Bandy ALE, McGroarty KK. Ethanol consumption by C57BL/6 mice: Influence of gender and procedural variables. *Alcohol*. 1999. [https://doi.org/10.1016/S0741-8329\(98\)00055-X](https://doi.org/10.1016/S0741-8329(98)00055-X).
31. Jackson LR, Robinson TE, Becker JB. Sex differences and hormonal influences on acquisition of cocaine self-administration in rats. *Neuropsychopharmacology*. 2006. <https://doi.org/10.1038/sj.npp.1300778>.
32. Simon NW, Gilbert RJ, Mayse JD, Bizon JL, Setlow B. Balancing risk and reward: A rat model of risky decision making. *Neuropsychopharmacology*. 2009. <https://doi.org/10.1038/npp.2009.48>.
33. Lobo MK, Covington HE, Chaudhury D, Friedman AK, Sun HS, Damez-Werno D, et al. Cell type—specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* (80). 2010. <https://doi.org/10.1126/science.1188472>.
34. Soares-Cunha C, Coimbra B, Sousa N, Rodrigues AJ. Reappraising striatal D1- and D2-neurons in reward and aversion. *Neurosci Biobehav Rev*. 2016;68:370–86.
35. Harkin T, Snowe OJ, Mikulski BA, Waxman HA. US General Accounting Office: Most Drugs Withdrawn in Recent Years Had Greater Health Risks for Women. *US Gen Account Off*. 2001;01-286R.
36. Beery AK, Zucker I. Sex bias in neuroscience and biomedical research. *Neurosci Biobehav Rev*. 2011;35:565–72.
37. Lee SK. Sex as an important biological variable in biomedical research. *BMB Rep*. 2018;51:167–73.