

# Dissociable Contributions of Basolateral Amygdala and Ventrolateral Orbitofrontal Cortex to Flexible Learning Under Uncertainty

C. G. Aguirre,<sup>1\*</sup> J. H. Woo,<sup>2\*</sup> J. L. Romero-Sosa,<sup>1</sup> Z. M. Rivera,<sup>1</sup> A. N. Tejada,<sup>1</sup> J. J. Munier,<sup>3</sup> J. Perez,<sup>1</sup> M. Goldfarb,<sup>1</sup> K. Das,<sup>1</sup> M. Gomez,<sup>1</sup> T. Ye,<sup>1</sup> J. Pannu,<sup>3</sup> K. Evans,<sup>1</sup> P. R. O'Neill,<sup>4</sup> I. Spigelman,<sup>3</sup> A. Soltani,<sup>2</sup> and A. Izquierdo<sup>1</sup>

<sup>1</sup>Department of Psychology, University of California, Los Angeles, California 90095, <sup>2</sup>Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755, <sup>3</sup>Section of Biosystems and Function, School of Dentistry, University of California, Los Angeles, California 90095, and <sup>4</sup>Shirley and Stefan Hatos Center for Neuropharmacology, Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, California 90095

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 the highly interconnected orbitofrontal cortex (OFC) and basolateral amygdala (BLA), and is often impaired in various neuropsychiatric and substance use disorders. However, the unique contributions of these regions to stimulus- and action-based reversal learning have not been systematically compared using a chemogenetic approach particularly before and after the first reversal that introduces new uncertainty. Here, we examined the roles of ventrolateral OFC (vOFC) and BLA during reversal learning. Male and female rats were prepared with inhibitory designer receptors exclusively activated by designer drugs targeting projection neurons 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 different reward probabilities. Using a counterbalanced within-subject design, we inhibited these regions prior to reversal sessions. We assessed initial and pre-/post-reversal changes in performance to measure learning and adjustments to reversals, respectively. We found that inhibition of the ventrolateral orbitofrontal cortex (vOFC), but not BLA, eliminated adjustments to stimulus-based reversals. Inhibition of BLA, but not vOFC, selectively impaired action-based probabilistic reversal learning, leaving deterministic reversal learning intact. vOFC exhibited a sex-dependent role in early adjustment to action-based reversals, but not in overall learning. These results reveal dissociable roles for BLA and vOFC in flexible learning and highlight a more crucial role for BLA in learning meaningful changes in the reward environment.

**Key words:** action learning; deterministic; DREADDs; probabilistic; reward learning; stimulus learning

## Significance Statement

Inflexible learning is a feature of several neuropsychiatric disorders. We investigated how the ventrolateral orbitofrontal cortex (vOFC) and basolateral amygdala (BLA) are involved in learning stimuli or actions under reinforcement uncertainty. Following chemogenetic inhibition of these regions in both males and females, we measured learning and adjustments to deterministic and probabilistic reversals. For action learning, BLA, but not vOFC, is needed for probabilistic reversal learning. However, BLA is not necessary for initial probabilistic learning or retention, indicating a critical role in learning of unexpected changes. For stimulus learning, vOFC, but not BLA, is required for adjustments to reversals, particularly in females. These findings provide insight into the complementary cortico-amygdalar substrates of learning under different forms of uncertainty.

Received April 2, 2023; revised Oct. 16, 2023; accepted Oct. 17, 2023.

Author contributions: C.G.A. and A.I. designed research; C.G.A., J.L.R.-S., Z.M.R., A.N.T., J.J.M., J.P., M.G., K.D., M.G., T.Y., J.P., and K.E. performed research; C.G.A., J.H.W., J.L.R.-S., J.J.M., P.R.O., A.S., and A.I. analyzed data; C.G.A., J.H.W., and A.I. wrote the paper.

This work was supported by UCLA's Division of Life Sciences Retention Fund (A.I.); National Institutes of Health Grants R01 DA047870 (A.I. and A.S.), R21 MH122800 (A.I. and Blair), R01AA024527 (I.S.), K01 DA042219 (P.O.), and F31 AA028183 (J.M.); the NSF GRFP, Cota-Robles Fellowship, and Charles E. and Sue K. Young Fellowship (C.A.); Ursula Mandel Fellowship and Graduate Research Mentorship Award (J.R.); and the Training Program in Neurotechnology Translation T32 NS115753 (T.Y.). We acknowledge

the Staglin Center for Brain and Behavioral Health for additional support related to fluorescence microscopy. We thank P. Ganupuru for his assistance with brain collection. We also thank the NIDA Drug Supply program for the supply of clozapine-N-oxide.

\*C.G.A. and J.H.W. are the co-first authors.

The authors declare no competing financial interests.

Correspondence should be addressed to Claudia Aguirre at cgaguirre@ucla.edu or Alicia Izquierdo at aizquie@psych.ucla.edu

<https://doi.org/10.1523/JNEUROSCI.0622-23.2023>

Copyright © 2024 the authors

## Introduction

Reversal learning, impacted in various neuropsychiatric conditions, measures subjects' ability to form flexible associations between stimuli and/or actions with outcomes (Schoenbaum et al., 2003; Izquierdo et al., 2013; Dalton et al., 2016). 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; Winstanley and Floresco, 2016; Soltani and Izquierdo, 2019). For example, after the experience of the first reversal, all other reversals are expected to some extent (Jang et al., 2015). Additionally, unexpected uncertainty can be introduced by changes in reward probabilities, after taking the baseline, expected uncertainty into account.

The BLA is an area of interest in reversal learning due to its involvement in value updating (Tye and Janak, 2007; Janak and Tye, 2015; Wassum and Izquierdo, 2015; Groman et al., 2019) and the encoding of both stimulus–outcome and action–outcome associations typically probed in Pavlovian-to-instrumental tasks (Corbit and 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 (Schoenbaum et al., 2003; Churchwell et al., 2009; Groman et al., 2019), impaired learning from positive feedback (Costa et al., 2016; Groman et al., 2019), enhanced learning from negative feedback (Rudebeck and Murray, 2008; Izquierdo et al., 2013; 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 versus 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 versus 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 vs all subsequent reversals) has also not been extensively studied using a chemogenetic approach.

In parallel, studies with manipulations in rat OFC in reversal learning have included targeting the entire ventral surface (Izquierdo et al., 2013; Izquierdo, 2017) or more recent systematic comparisons of medial versus lateral OFC (Hervig et al., 2020; Verharen et al., 2020). Here, we examined the role of vOFC, a subregion not as often probed in reward learning as medial and more (dorso)lateral OFC [compare Zimmermann et al. (2018)] but also densely interconnected with BLA (Barreiros et al., 2021a,b). 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 adjustments to reversals. We found that vOFC, but not BLA, inhibition impaired adjustments to deterministic and probabilistic reversals. Conversely, BLA, but not vOFC, inhibition resulted in significantly slower action-based probabilistic, but not deterministic, reversal learning. Fitting choice data with reinforcement learning (RL) models indicated that action-based reversal learning deficits were mediated by a larger memory decay for the unchosen option following vOFC inhibition and diminished exploration after reversal following

BLA inhibition. These results suggest dissociable roles for BLA in flexible learning under uncertainty, and vOFC in adjustments to reversals, more generally.

## Materials and Methods

### Subjects

Animals for behavioral experiments were adult ( $N=70$ , 33 females; 66 used for behavioral study and 4 males for ex vivo imaging) Long-Evans rats (Charles River Laboratories) with average age postnatal day (PND) 65 at the start of experiments, with a 280 g body weight minimum for males and 240 g 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-d acclimation period during which they were pair-housed and given food and water ad libitum. During that time, they remained in their home cage with no experimenter interference. Following this 3-d 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 6 A.M.). Animals were then surgicized and tested on discrimination and reversal learning 1-week postsurgery. At the point of reversal, they were beyond the 3-week expression time for designer receptors exclusively activated by designer drugs (DREADDs).

A separate group of Long-Evans rats ( $N=4$ , all males) was used for validation of the effectiveness of DREADDs in slides of BLA and vOFC, using ex vivo calcium imaging procedures. 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 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 home cages for 4 weeks prior to testing while the inhibitory hM4Di DREADDs expressed in BLA ( $n=31$ , 16 females), vOFC ( $n=19$ , 10 females), or enhanced green fluorescent protein (eGFP) control virus ( $n=16$ , 7 females) in these regions. In rats tested on behavior, 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  $\mu$ l, AP =  $-2.5$ ; ML =  $\pm 5$ ; DV =  $-7.8$  and 0.2  $\mu$ l, AP =  $-2.5$ ; ML =  $\pm 5$ ; DV =  $-8.1$ , from bregma at a rate of 0.1  $\mu$ l/min; AAV8-CaMKIIa-hM4D(Gi)-mCherry, Addgene, viral prep #50477-AAV8]. In other animals, this same virus (AAV8-CaMKIIa-hM4Di-mCherry, Addgene) was bilaterally infused into two sites in vOFC (0.2  $\mu$ l, AP =  $+3.7$ ; ML =  $\pm 2.5$ ; DV =  $-4.6$  and 0.15  $\mu$ l, AP = 4; ML =  $\pm 2.5$ ; DV =  $-4.4$ , from bregma at a rate of 0.1  $\mu$ l/min). A virus lacking the hM4Di DREADD gene and only containing the green fluorescent tag eGFP (AAV8-CaMKIIa-eGFP, Addgene) was also infused bilaterally into either BLA ( $n=7$ ), vOFC ( $n=5$ ), or anterior cingulate cortex [( $n=5$ ); 0.3  $\mu$ l, AP =  $+3.7$ ; ML =  $\pm 2.5$ ; DV =  $-4.6$ , rate of 0.1  $\mu$ l/min] as null virus controls. Our vOFC targeting is most similar to infusion sites reported previously by Dalton et al. (2016) constituting lateral as well as ventral OFC and 0.7 mm more medial than others (Costa et al., 2023). In rats used for ex vivo calcium imaging, the same target 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<sub>2</sub>) anesthesia prior to any behavioral testing experience. Before surgeries were completed, all animals were subcutaneously administered 5 mg/kg carprofen

(NADA #141–199, Pfizer, Drug Labeler Code 000069) and 1 cc saline. After being placed in the stereotaxic apparatus (David Kopf; model 306041), the scalp was incised and retracted. The skull was leveled with a  $\pm 0.3$  mm tolerance on the AP to ensure that bregma and lambda were in the same horizontal plane. Small burr holes were drilled in the skull above the infusion target. The virus was bilaterally infused at a rate of 0.01  $\mu$ l 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 (0.8 ml, 390 mg/ml pentobarbital, 50 mg/ml phenytoin; Virbac) were transcardially perfused, and their brains were removed for histological processing. Brains were fixed in 10% buffered formalin acetate for 24 h followed by 30% sucrose for 5 d. To visualize hM4Di-mCherry and eGFP expression in BLA or vOFC cell bodies, free-floating 40  $\mu$ m coronal sections were mounted onto slides and coverslipped with mounting medium for DAPI. Slices were visualized using a BZ-X710 microscope (Keyence) and analyzed with BZ-X Viewer and analysis software.

Reconstructions of viral expressions of hM4Di (magenta) and eGFP (green) across the AP plane (Fig. 1*B,E*) were conducted using Photoshop and Illustrator (Adobe). Two independent raters blind to the condition then used ImageJ (US National Institutes of Health) to trace and quantify pixels at AP +3.7 (vOFC) and AP –2.8 (BLA) for each animal. Three measures were obtained per hemisphere and rater measurements were significantly correlated (Pearson correlation:  $r = 0.54$ ,  $p = 4.51e^{-04}$ ). There were no differences in expression level between males and females for pixel count reconstructions [ $F_{(1,36)} = 2.59$ ,  $p = 0.12$ ]. Only subjects with bilateral expression were included in behavioral analyses (four vOFC and eight BLA hM4Di rats were excluded due to unilateral expression).

### Food restriction

Five days prior to any behavioral testing, rats were placed on food restriction with females on average maintained on 10–12 g/d and males given 12–14 g/d 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 vOFC or BLA was achieved by systemic administration of clozapine-*N*-oxide, CNO (3 mg/kg, i.p., 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 before learning, 30 min prior to behavioral testing. We followed previous work on the timing and dose of systemic CNO (Stolyarova et al., 2019; Hart et al., 2020) and considered the long duration of test sessions. To control for nonspecific effects of injections and handling stress, we also injected animals with saline vehicle (VEH). For reversal learning, 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 a counterbalanced order. Thus, for drug administration, if a rat received CNO on the first reversal (R1), it was administered VEH on the second reversal (R2), CNO on the third reversal (R3), and VEH on the fourth reversal (R4) or vice versa: VEH on R1, CNO on R2, VEH on R3, and CNO on R4.

### 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 (Lafayette Instrument).

The pretraining protocol, adapted from established procedures (Stolyarova and Izquierdo, 2017), consisted of a series of phases: habituation, initiation touch to center training (ITCT), and immediate reward

training (IMT), designed to train rats to nose poke, 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 the collection of all five sucrose pellets. For ITCT, the criterion for 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 2 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 (Figs. 2*A,B*, 6*A,B*). A subset of animals was tested first on the action-based task (13 vOFC hM4Di, 10 BLA hM4Di), while others were tested on the stimulus-based task first (9 vOFC hM4Di, 7 BLA hM4Di, 16 eGFP). Three rats in the vOFC group 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 (Fig. 2*A*). 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 stimuli (i.e., fan or marble) on the left and right side of the screen (displayed for 60 s). Rats could nose poke either the spatial side rewarded with one sucrose pellet [the better option,  $p_R(B) = 1.0$ ] or the spatial side that went unrewarded [the worse option,  $p_R(W) = 0.0$ ]. Thus, rats were required to ignore the properties of the stimuli and determine the better-rewarded side. If a side was not selected, it was scored as a choice omission, and a 10 s inter-trial interval (ITI) ensued. If a trial was not rewarded, a 5 s time-out would occur, followed by a 10 s ITI. If a trial was rewarded, a 10 s ITI would occur after the reward was collected. 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 2 consecutive days. After reaching the criterion for the discrimination phase, the rats were advanced to the reversal phase beginning in 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 CNO or VEH 30 min prior to each reversal testing session. The side previously associated with the  $p_R(B) = 1.0$  probability was now associated with a  $p_R(W) = 0.0$  probability of being rewarded 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 in the next session. Rats that had previously received VEH during the first reversal would now receive CNO injections and vice versa.

After completing both deterministic reversal learning phases, rats advanced to the first probabilistic reversal learning phase (i.e., R3). Rats underwent the same injection procedure as the prior reversals. However, the spatial side (i.e., left or right), previously associated with  $p_R(B) = 1.0$ , was now associated with a  $p_R(W) = 0.1$  probability of being rewarded, whereas the spatial side, previously associated with  $p_R(W) = 0.0$  probability, was now associated with  $p_R(B) = 0.9$ . 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., R3), rats were advanced to the second probabilistic reversal phase (i.e., 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 (Fig. 6A). Stimuli were displayed for 60 s each, randomly assigned as the better or worse stimulus:  $p_R(B) = 1.0$  or  $p_R(W) = 0.0$ . If a trial was not initiated within 40 s, it was scored as an initiation omission. If a stimulus was not selected, 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 occur, 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 2 consecutive days. After reaching the criterion for the discrimination phase or if rats were unable to achieve the criterion after 10 d, rats were advanced to the reversal phase beginning in the next session. Animals were not administered 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., 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 VEH 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 the criterion after 10 d, the rats were advanced to the second deterministic reversal phase (i.e., 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., R3). The injection procedure remained the same as prior reversals. However, the visual stimulus previously associated with  $p_R(B) = 1.0$  would now be associated with  $p_R(W) = 0.1$ , whereas the stimulus previously associated with  $p_R(W) = 0.0$  would now be associated with  $p_R(B) = 0.9$ . 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 the criterion after 10 d, the rats were advanced to the second probabilistic reversal phase (i.e., 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.

**Action-based probabilistic learning and retention.** We assessed initial probabilistic learning, using the same criterion as above, in a separate group of experimentally naive animals expressing hM4Di in BLA ( $N = 14$ , 7 females; 2 animals were excluded due to unilateral expression). Before a learning session, rats were injected intraperitoneally with either 3 mg/kg of CNO ( $n = 10$ ) or VEH ( $n = 2$ ) 30 min prior to testing. One side of the touchscreen was associated with a better probability of reward [ $p_R(B) = 0.9$ ], and the other side was associated with a worse probability of reward [ $p_R(W) = 0.1$ ]. Animals were then assessed for their retention of this association in the next session.

### Ex vivo calcium imaging

In  $N = 4$  animals (all males), following >3 weeks following stereotaxic viral injections, rats ( $n = 1$  rat/brain region/virus combination;  $n = 2$ –5 slices/rat) were deeply anesthetized with isoflurane (Patterson Veterinary) and decapitated, and brains were submerged in ice-cold oxygenated (95/5%  $O_2/CO_2$ ) slicing artificial cerebrospinal fluid (ACSF) containing the following (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 vOFC or BLA slices (300  $\mu$ M thick) were obtained (VT1200s, Leica) and transferred to room temperature normal ACSF containing the following (in mM): 125 NaCl, 2.5 KCl, 1.25  $NaH_2PO_4$ , 2  $CaCl_2$ , 2  $MgCl_2$ , 26  $NaHCO_3$ , and 10 glucose, pH adjusted to 7.3 with KOH and allowed equilibrate for >1 h prior to transfer into a perfusion 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 4.0 v3). Anatomical regions in brain sections for  $Ca^{2+}$  imaging were first identified by brightfield imaging with 780 nm LED (Scientifica) illumination.  $Ca^{2+}$  imaging was performed using a 40 $\times$ , 0.80NA water immersion objective (Olympus), continuous 470 nm LED illumination (Thorlabs), and a filter cube suitable for GCaMP6f imaging: excitation, BrightLine 466/40; dichroic, Semrock FF495-Di03; emission, BrightLine 525/50. Slices were housed on poly-D-lysine coverslips attached to the RC-26G chamber (Warner Instruments), which was modified with platinum wires to apply electric field stimulation. Images were acquired continually with 20 ms exposure time. Electric field stimulation was applied at 110 mV (train pulse every 5 s). The temperature of ACSF during the recorded sessions was held at 28°C to minimize bubble formation.

### Calcium data extraction

Prior to imaging sessions, 40 $\times$  images of red and green fluorescence were captured and subsequently overlaid for post hoc genotyping of individual cells (GCaMP6f<sup>+</sup>, GCaMP6f<sup>+</sup>/hM4Di-mCherry<sup>+</sup>, or GCaMP6f/mCherry<sup>+</sup>). Blinded scorers semi-manually curated ROIs using Python-based suite2P software (Pachitariu et al., 2017). ROI fluorescence was subtracted from the annular surround fluorescence, low-pass filtered, and transformed to  $dF/F_0$  as previously described (Asrican and 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 9,000 to eliminate negative changes. The area under the curve and event frequency of each cell were 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; Fig. 1C,F).

### Data analyses

MATLAB (MathWorks, Version R2021a) was used for all statistical analyses and figure preparation. Data were analyzed with a series of mixed-effect general linear models (GLMs, *fitglm*) 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-effect 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 an asymptote at 5 d for stimulus-based reversal learning, only the first 5 d 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 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 and +3 sessions or –100 trials and +100 trials 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) on pixel counts obtained via ImageJ. Pearson correlation coefficients (*corrcoef*) were analyzed for inter-rater reliability of viral expression quantification. Wilcoxon rank sum tests were used to compare RL parameters between groups.

Dependent measures for learning included the probability of choosing the correct or better option, probability of win–stay, and probability of lose–switch. The probability of win–stay and lose–switch adaptive strategies was 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 the number of comparisons.

To analyze ex vivo calcium imaging data, two-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 GCaMP6f and GCaMP6f + hM4Di-mCherry in each brain region for the experimental group. Tests corrected for the number of comparisons were conducted for interactions.

### RL models

To capture differences between groups in learning and choice behavior during the action-based task, we utilized two conventional 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:

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

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

The second model, referred to as RL<sub>decay</sub>, 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:

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

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

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

where  $i$  and  $j$  correspond to two alternative options (i.e., left and right for an action-based task) and  $\beta$  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  $\alpha$  and decay rate  $\gamma_d$  were bounded between 0 and 1, and  $\beta$  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 100 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<sub>decay</sub>), 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 the Akaike information criterion (AIC) and Bayesian information criterion (BIC) for each session as follows:

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

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

where  $k$  is the number of parameters in the model (two for RL and three for RL<sub>decay</sub>) and  $n$  is the number of choice trials in the session.

## Results

### Ex vivo calcium imaging in slices

We performed ex vivo  $Ca^{2+}$  imaging to confirm the selective action on CaMKII<sup>+</sup> neuronal excitability in vOFC and BLA in

rats expressing hM4Di DREADD versus controls expressing mCherry. In BLA, there was no significant effect of CNO (10  $\mu$ M) on  $Ca^{2+}$  events for neurons expressing GCaMP6f or GCaMP6f + mCherry (Fig. 1C). A two-way ANOVA resulted in a significant drug  $\times$  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 vOFC, there was also no significant effect of CNO (10  $\mu$ M) on  $Ca^{2+}$  events for neurons expressing GCaMP6f or GCaMP6f + mCherry (Fig. 1F). However, in CaMKII<sup>+</sup> vOFC neurons expressing GCaMP6f + hM4Di, there was a decrease in the frequency of  $Ca^{2+}$  events during CNO application. A two-way ANOVA revealed a significant drug  $\times$  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

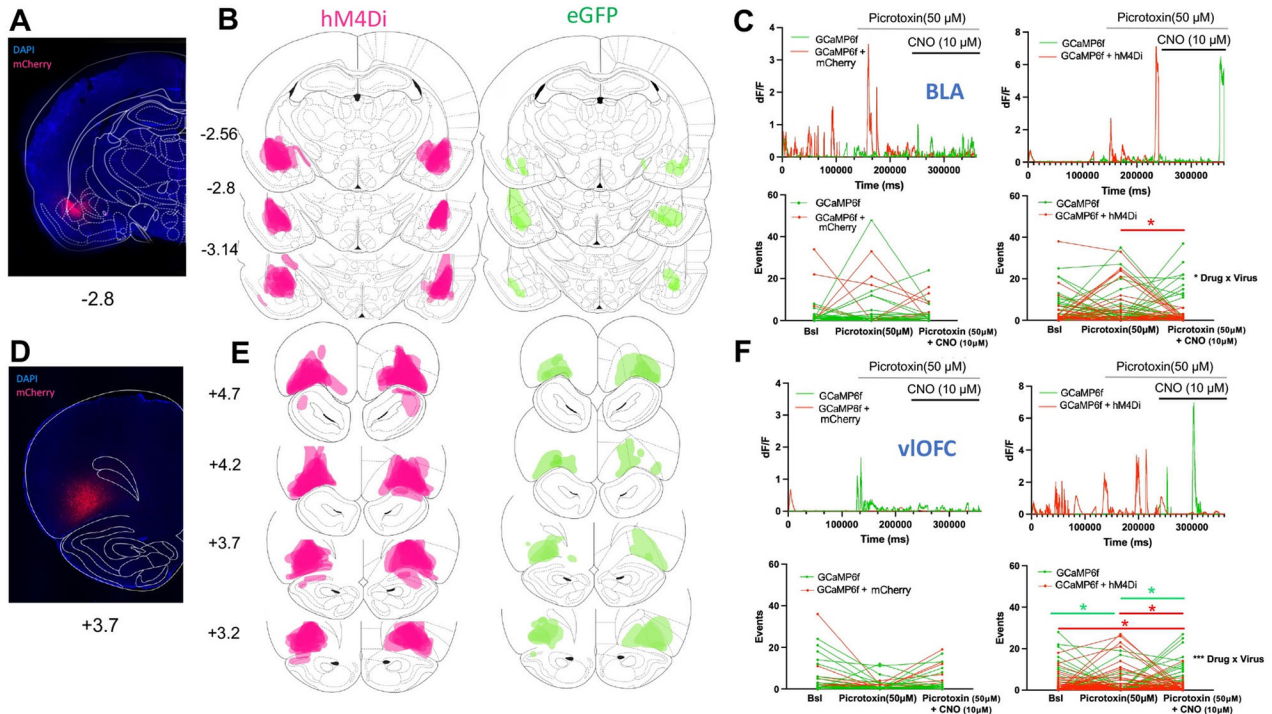
Mixed-effect GLMs for the discrimination learning phase were conducted for each task separately to establish if there were baseline differences in learning measures between animals infused with the eGFP virus in different brain regions. There were no differences between the eGFP groups by target region on learning (i.e., the probability of choosing the correct side) across trials in the action-based task ( $\beta_{\text{region}} = -0.13$ ,  $p = 0.47$ ), as well as no differences in learning (i.e., probability of choosing the correct visual stimulus) across sessions in the stimulus-based task ( $\beta_{\text{region}} = 0.10$ ,  $p = 0.09$ ). Thus, animals' data were collapsed into a single eGFP virus group for subsequent analyses.

### Discrimination learning: hM4Di versus eGFP

For the action-based task, there were no significant effects of virus or virus interactions for vOFC versus eGFP on *probability correct* ( $\beta_{\text{virus}} = -0.002$ ,  $p = 0.96$ ), with similar findings for the comparison of BLA versus eGFP ( $\beta_{\text{virus}} = -0.02$ ,  $p = 0.85$ ; Fig. 2C). All animals met criterion very quickly ( $\sim 2$  d); thus, we compared trials to reach 75% criterion (i.e., probability of choosing the correct side). Both hM4Di virus groups performed comparably [ $M \pm \text{SEM}$ : vOFC hM4Di (69.6  $\pm$  15.4), BLA hM4Di (80.9  $\pm$  21.5)], whereas the eGFP group met criterion within fewer trials (59.5  $\pm$  18.6), but the difference was not statistically significant [vOFC hM4Di vs eGFP:  $\beta_{\text{virus}} = 33.2$ ,  $p = 0.34$ ; BLA hM4Di vs eGFP:  $\beta_{\text{virus}} = 51.3$ ,  $p = 0.21$ ].

For the stimulus-based task, there were also no significant effects of virus or virus interactions for either vOFC versus eGFP on *probability correct* ( $\beta_{\text{virus}} = -0.09$ ,  $p = 0.10$ ) or for BLA versus eGFP ( $\beta_{\text{virus}} = -0.06$ ,  $p = 0.20$ ; Fig. 6C). The animals on average took approximately  $\sim 6$  d to meet criterion regardless of virus group [ $M \pm \text{SEM}$ : vOFC hM4Di (6.1  $\pm$  0.7), BLA hM4Di (6.5  $\pm$  1.2), eGFP (6.9  $\pm$  0.7)].

Given the poorer learning in the stimulus-based task, we evaluated whether this was due to the order of task administered [i.e., stimulus  $\rightarrow$  action or action  $\rightarrow$  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 ( $\beta_{\text{order}} = 0.03$ ,  $p = 0.33$ ) but a significant task order  $\times$  session interaction ( $\beta_{\text{order} \times \text{session}} = -0.03$ ,  $p = 0.002$ ). Thus, subsequent analyses were conducted with task order analyzed separately by session, which revealed that animals administered the action  $\rightarrow$  stimulus task order exhibited poorer learning across sessions ( $\beta_{\text{session}} = 0.01$ ,  $p = 0.05$ ).



**Figure 1.** Bilateral targeting of BLA and vIOFC and confirmation of effective DREADD inhibition using ex vivo  $\text{Ca}^{2+}$  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 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 (10  $\mu\text{M}$ ) in the presence of picrotoxin (50  $\mu\text{M}$ ) had no effect on the frequency of elicited  $\text{Ca}^{2+}$  events (top left and right, example of single-cell traces; bottom left,  $\text{Ca}^{2+}$  event changes, each line is a cell). In BLA neurons that expressed hM4Di, there was a reduction in the frequency of elicited  $\text{Ca}^{2+}$  events during CNO application (bottom right). **D**, Photomicrograph of hM4Di-mCherry DREADDs expression in vIOFC. Numerals indicate mm anterior to Bregma. **E**, Reconstruction of viral expression of hM4Di (magenta) and eGFP (green). The more intense colors represent regions where expression overlapped the most across animals. **F**, In vIOFC neurons expressing GCaMP6f and GCaMP6f + mCherry, application of CNO (10  $\mu\text{M}$ ) had no effect on the frequency of elicited  $\text{Ca}^{2+}$  events (top left and right, example of single-cell traces; bottom left,  $\text{Ca}^{2+}$  event changes; each line is a cell). In vIOFC neurons that expressed hM4Di, there was a reduction in the frequency of  $\text{Ca}^{2+}$  events during CNO application (bottom right).  $n = 2\text{--}5$  slices/rat, two-way ANOVA and multiple comparison tests. \*\*\* $p < 0.001$ , \* $p < 0.05$ .

compared to those administered the stimulus  $\rightarrow$  action task order ( $\beta_{\text{session}} = 0.04$ ,  $p < 0.0001$ ). To further address the possibility that any effects on reversal learning could be due to task order instead of DREADD inhibition, we conducted a separate assessment of learning in animals administered the action-based task first (see below).

### Accuracy across trials and sessions: reversal learning

#### Action-based reversal learning

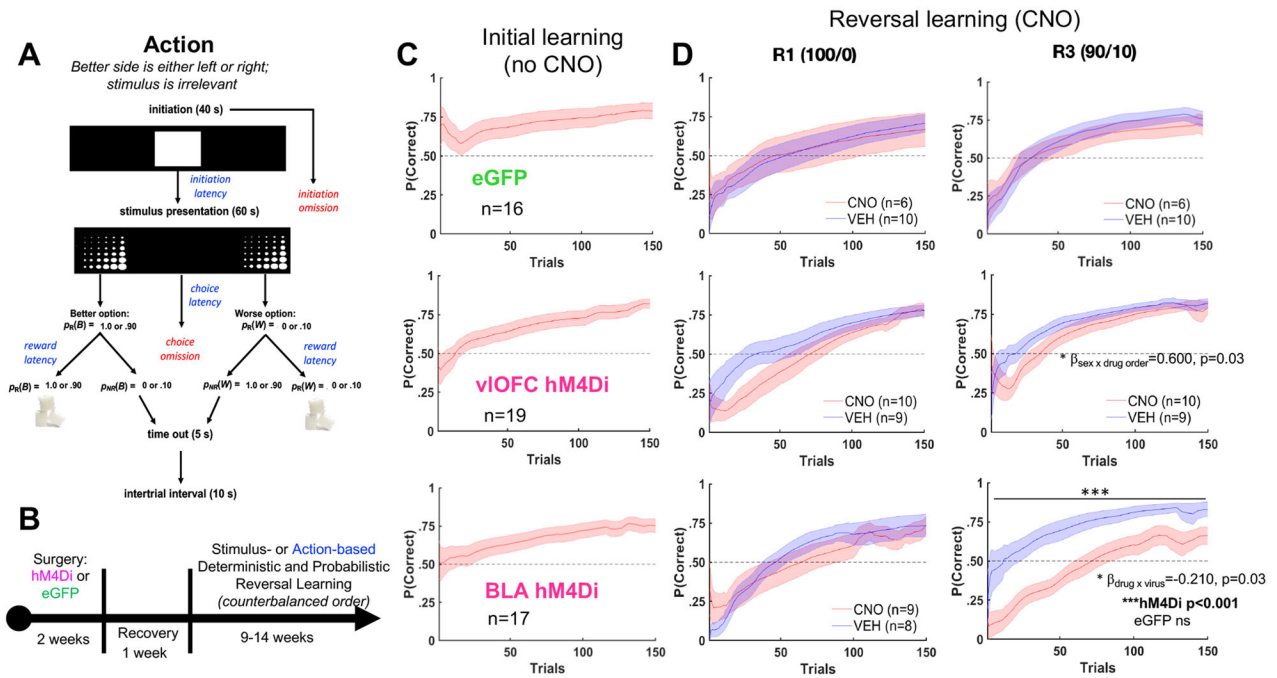
Mixed-effect GLMs were used to analyze *probability correct*, our primary measure of accuracy, with drug order, virus, and sex as between-subject factors; trial number, reversal number, and drug as within-subject factors; and individual rat as a random factor. GLMs were conducted separately by target region (BLA vs eGFP and vIOFC vs eGFP), using the following formula for the full model:  $\gamma \sim [1 + \text{trial number} + \text{reversal number} \times \text{virus} \times \text{drug} \times \text{drug order} \times \text{sex} + (1 + \text{trial number} + \text{reversal number} \times \text{drug}) \text{rat}]$ .

Starting with the full model above for the comparison of BLA with eGFP, we found that trial number was a significant predictor of *probability correct* ( $\beta_{\text{trial number}} = 0.003$ ,  $p = 3.26e^{-81}$ ). We also observed a sex  $\times$  drug order ( $\beta_{\text{sex} \times \text{drug order}} = 0.63$ ,  $p = 0.01$ ) and sex  $\times$  virus  $\times$  drug order ( $\beta_{\text{sex} \times \text{virus} \times \text{drug order}} = -0.72$ ,  $p = 0.046$ ) interaction. To further probe “first reversals,” we included only R1 and R3 in the above model. Accordingly, drug orders did not vary across R1 and R3. In the analysis of first experiences with deterministic R1 and probabilistic R3, the trial number was also a significant predictor of accuracy ( $\beta_{\text{trial number}} = 0.003$ ,

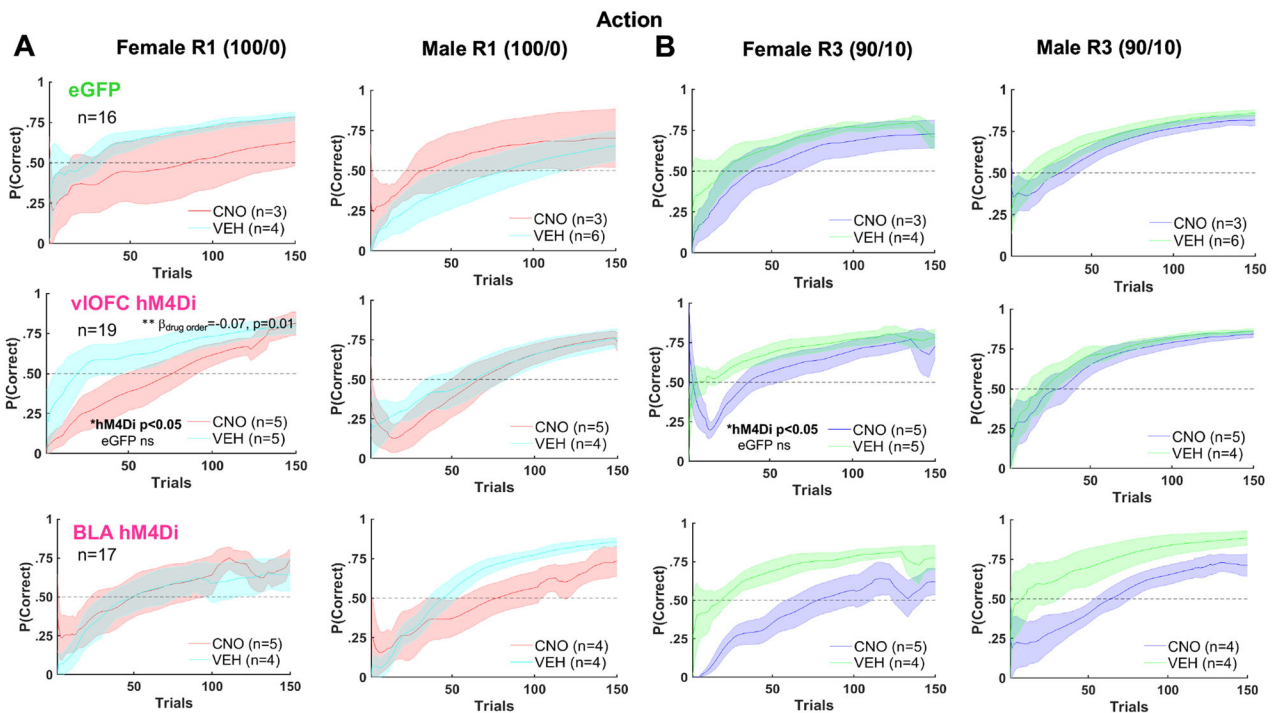
$p = 2.59e^{-49}$ ), along with a virus  $\times$  drug  $\times$  reversal number interaction ( $\beta_{\text{virus} \times \text{drug} \times \text{reversal number}} = -0.19$ ,  $p = 0.04$ ), which justified further analysis of R1 and R3, separately. Follow-up GLMs were conducted for R1 and R3 *probability correct*, with the following formula:  $\gamma \sim [1 + \text{virus} \times \text{drug} + (1 | \text{rat})]$ . A significant virus  $\times$  drug interaction was obtained only for R3 ( $\beta_{\text{virus} \times \text{drug}} = -0.22$ ,  $p = 0.03$ ; Fig. 2D). Bonferroni-corrected post hoc comparisons revealed an effect of CNO in hM4Di ( $p = 2.33e^{-04}$ ), not in eGFP ( $p = 1.0$ ), with both males and females exhibiting impaired learning of probabilistic R3 following BLA inhibition (Fig. 3B). Importantly, stimulus-naïve animals also demonstrated impaired probabilistic reversal learning ( $\beta_{\text{drug}} = -0.23$ ,  $p = 0.02$ ; Fig. 4A), indicating this slower learning was not due to previous training history.

For the comparison of vIOFC with eGFP, the trial number was a significant predictor of *probability correct* ( $\beta_{\text{trial number}} = 0.003$ ,  $p = 6.39e^{-75}$ ). We also observed a sex  $\times$  drug order ( $\beta_{\text{sex} \times \text{drug order}} = 0.60$ ,  $p = 0.03$ ) interaction (Fig. 2D). Follow-up GLMs were conducted for accuracy in females and males separately, with the following formula:  $\gamma \sim [1 + \text{drug order} + (1 | \text{rat})]$ . A significant drug order effect (CNO first > VEH first) was obtained for females ( $\beta_{\text{drug order}} = -0.07$ ,  $p = 0.01$ ) but not males (Fig. 3). Due to finding no interactions of reversal number with any other predictor, we were not justified to look further at individual reversals. Altogether, there was no effect of vIOFC inhibition on action-based reversal learning: all animals eventually reached asymptote and criterion over a comparable number of trials.

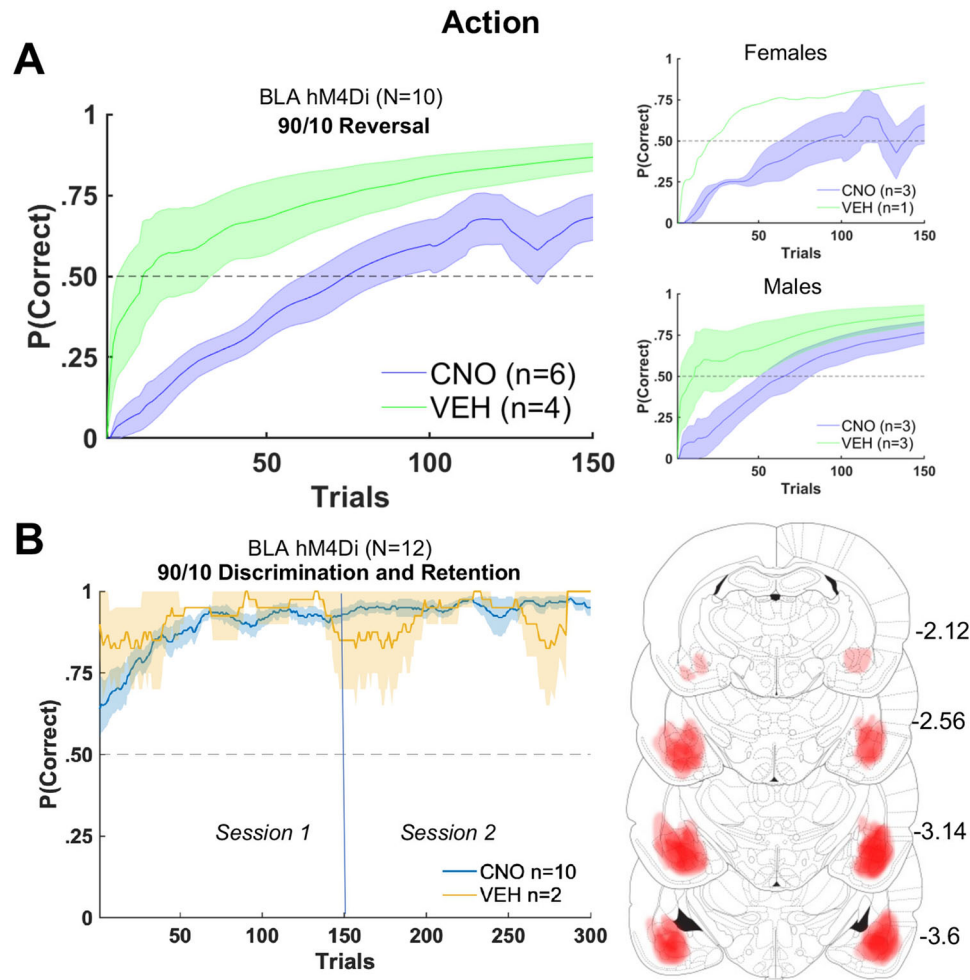




**Figure 2.** BLA inhibition impairs action-based probabilistic reversal learning, whereas vIOFC inhibition only slows early adjustment to reversals. **A, B**, Trial structure (**A**) and timeline (**B**) of the action-based task. Rats were first surgicized with either hm4Di DREADDs on a CaMKII promoter 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) reversal. Plots show cumulative P(Correct) for the first 150 trials with a sliding window of 10 trials. 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 inhibition on learning in R2 and R4 (not shown). There was no effect of CNO on learning in the eGFP group. \* $p < 0.05$  significant sex by drug order interaction (see Fig. 3 for learning curves plotted by sex). Bonferroni-corrected post hoc comparisons following mixed-effect GLM with sex as a covariate fixed factor wherein a drug  $\times$  virus interaction was found resulting in \*\*\* $p < 0.001$  effect of drug only in BLA hm4Di, not in eGFP. The impairment in probabilistic reversal learning following BLA inhibition was also evident in animals with no history of stimulus-based learning and was specific to reversal, not initial probabilistic learning (Fig. 4).



**Figure 3.** Female adjustment to reversals in the early phase was more affected by vIOFC inhibition than in males, whereas BLA's role in probabilistic reversal learning was not sex-dependent. Plotted is the accuracy for the first 150 trials with a sliding window of 10 trials. 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 vIOFC lowered P(Correct) in early first deterministic R1 and first probabilistic R3 in females but not males, whereas BLA inhibition attenuated probabilistic reversal learning in both males and females. Bonferroni-corrected post hoc comparisons upon GLM result of sex  $\times$  drug order interaction resulted in an effect of drug order only in vIOFC females, not in eGFP. The impairing effect of BLA inhibition on learning was not sex-dependent. There were no significant sex differences and no effect of CNO on learning in the eGFP group. \*\* $p = 0.01$ , \* $p < 0.05$ . The RL model fit to choice behavior also indicated that chemogenetic inhibition of vIOFC increased the decay rate (decreased memory) of the unchosen option ( $\gamma_d$ ) in female rats (Fig. 5).



**Figure 4.** BLA inhibition impairs action-based probabilistic reversal learning in stimulus-naïve animals, but does not affect initial action-based probabilistic discrimination or retention. **A**, Animals without any prior experience with stimulus-based learning exhibit impaired probabilistic reversal learning. **B**, There was no significant effect of BLA inhibition on initial probabilistic discrimination or next-day retention of the better-rewarded side/action. Bilateral expression of hM4Di DREADDs in BLA reconstructed from histology (right) in these animals.

Since OFC has been implicated in early reversal learning both classically and recently (Jones and Mishkin, 1972; Schoenbaum et al. 2007; Amodeo et al. 2017), we conducted additional analyses using the same full-model formula  $\gamma \sim [1 + \text{trial number} + \text{reversal number} \times \text{virus} \times \text{drug} \times \text{drug order} \times \text{sex} + (1 + \text{trial number} + \text{reversal number} \times \text{drug}) \text{ rat}]$ , but this time restricted to early trials (first 10) of each reversal. Within this early phase, there was a significant sex  $\times$  virus  $\times$  drug order ( $\beta_{\text{sex} \times \text{virus} \times \text{drug order}} = -0.94, p = 0.03$ ) and sex  $\times$  virus  $\times$  drug  $\times$  drug order interaction ( $\beta_{\text{sex} \times \text{virus} \times \text{drug} \times \text{drug order}} = 1.05, p = 0.04$ ). When sex was entered as a covariate, there was a marginally significant interaction of virus  $\times$  drug order for all reversals ( $\beta_{\text{virus} \times \text{drug order}} = -0.22, p = 0.056$ ), with a sex difference observed only in the hM4Di group ( $p = 0.03$ ), not in eGFP ( $p = 0.44$ ). Thus, early reversal adjustment in female rats was more adversely affected by previous OFC inhibition (i.e., if they received CNO first) than in males (Fig. 3).

#### Action-based probabilistic learning and retention

To further probe if the effect of BLA inhibition on probabilistic reversal learning was specific to reversal, and not initial learning, we administered CNO or VEH before initial 90/10 learning in a separate group of experimentally-naïve animals transfected with hM4Di in BLA. To analyze the effect of drug on probability

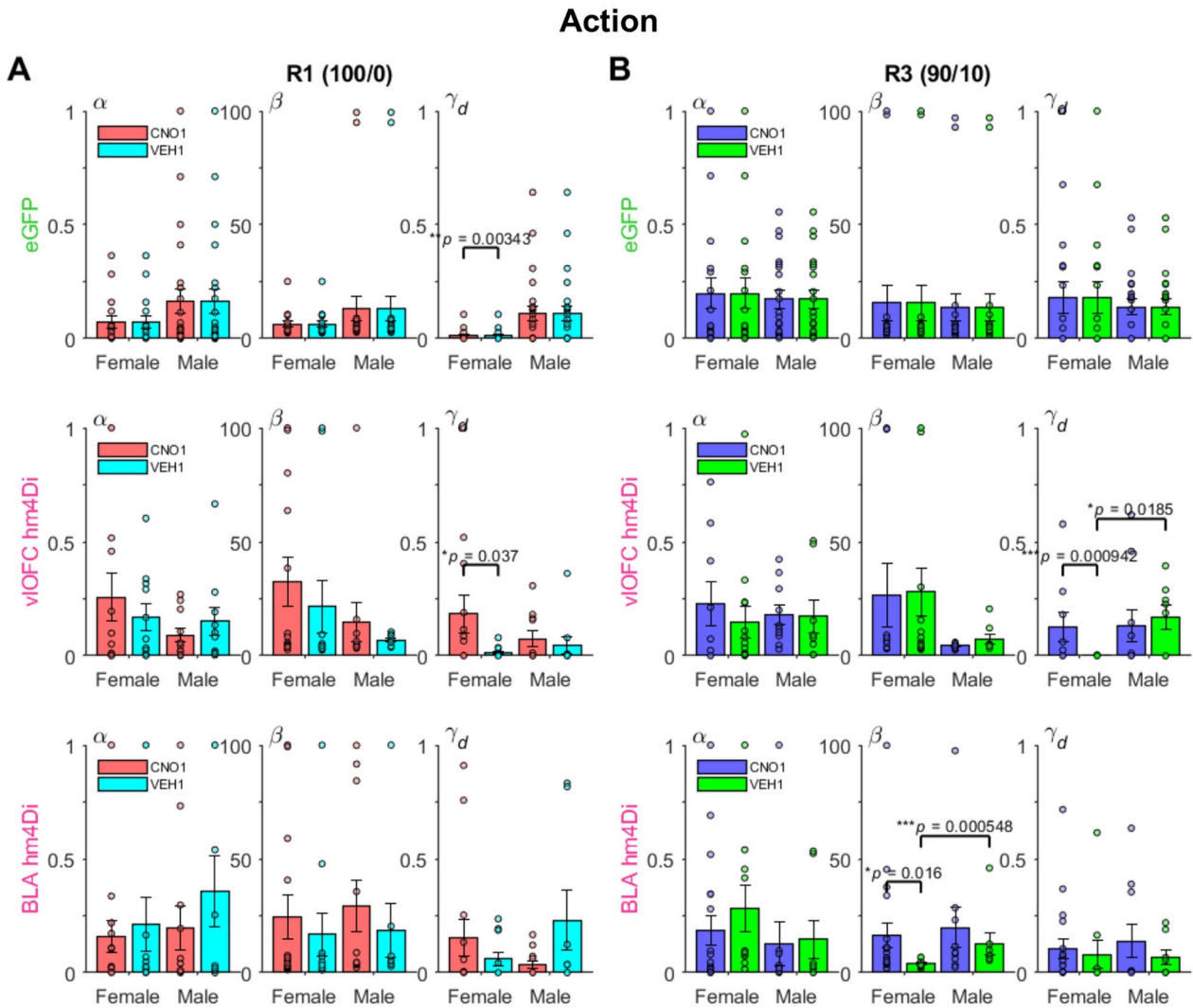
correct, we used the formula:  $\gamma \sim [1 + \text{trial number} + \text{drug} + (1 + \text{trial number}) \text{ rat}]$ . Only an effect of trial number was found ( $\beta_{\text{trial number}} = 0.001, p = 1.7e^{-04}$ ). There was no significant effect of drug on learning in either session 1 (initial learning) or session 2 (retention), with or without sex entered in the model as a covariate. Thus, all rats demonstrated learning and full retention on the next day (Fig. 4B).

#### Fitting choice behavior in R1 and R3 with RL models

To gain more insight into the effect of vlOFC and BLA inhibition on deterministic reversal learning (R1) and probabilistic reversal learning (R3) and their potential underlying mechanisms, we next compared the estimated model parameters from RL models (RL, RL<sub>decay</sub>). Comparing the goodness of fit between the two models, we found that the second model with the decay parameter (RL<sub>decay</sub>) better accounted for the animals' choice behavior as indicated by significantly lower AIC [paired sample *t* test;  $t_{(1334)} = 4.613, p = 4.34e^{-06}$ ]. In contrast, the overall mean BIC value was significantly lower for the first model [ $t_{(1334)} = -2.623, p = 8.00e^{-03}$ ]. We focused on the estimated parameters from the RL<sub>decay</sub> model only (Fig. 5), as AIC better distinguished the better-fitting model and allowed estimation of the additional  $\gamma_d$  parameter.

Comparison of the estimated parameters across groups revealed that female and male rats differed mainly in the decay





**Figure 5.** The RL model fit to choice behavior indicates differential effects on inverse temperature and decay parameters between male and female rats during deterministic and probabilistic action-based reversals. Plotted are estimated model parameters for RL1decay fitted to each session during deterministic (100/0) (**A**) and probabilistic (90/10) (**B**) reversals.  $\alpha$  = learning rate,  $\beta$  = inverse temperature,  $\gamma_d$  = 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. The female eGFP group exhibited overall lower  $\gamma_d$  than male eGFP rats during the first deterministic R1. Chemogenetic inhibition of vIOFC increased  $\gamma_d$  on both reversals (R1 and R3) for female rats, whereas inhibition of BLA increased  $\beta$  on the first probabilistic R3 for female rats. Only comparisons with  $*p < 0.05$ .

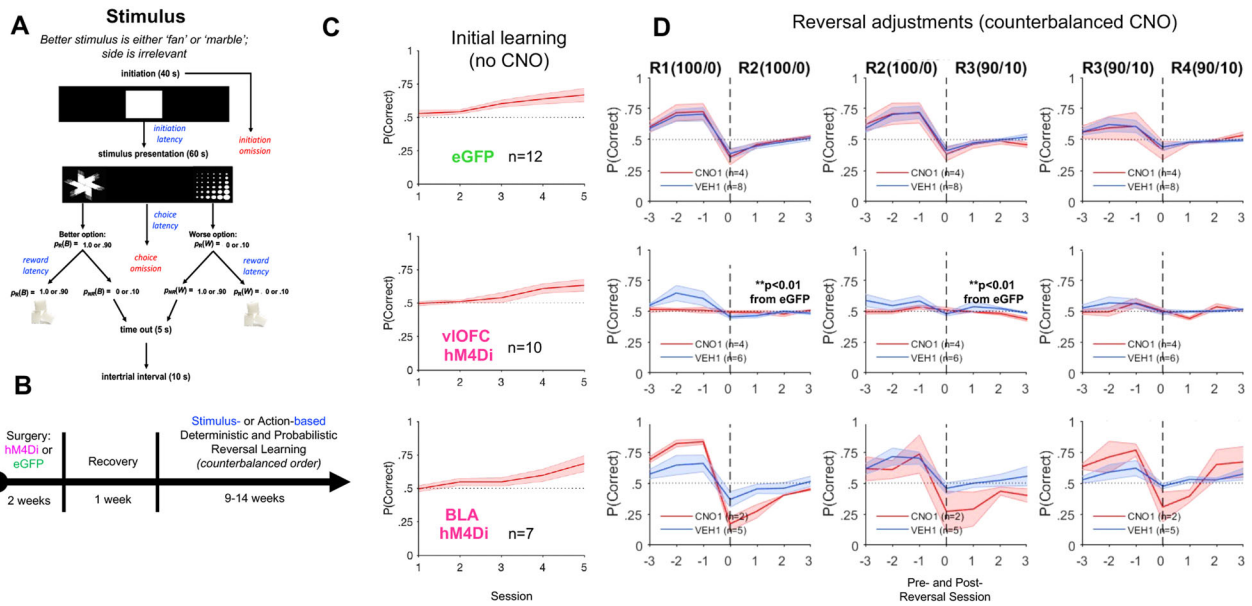
parameter  $\gamma_d$ , which governs the amount of passive decay or forgetting in the value estimate of the unchosen option. During the first deterministic reversal (R1, 100/0), eGFP females showed overall significantly lower values of  $\gamma_d$  than eGFP males (mean difference in  $\gamma_d = -0.0982$ ; Wilcoxon rank sum test,  $p = 3.40e^{-06}$ ), suggesting a different mechanism of adjustment to the reversal between females and males. While there was no clear evidence for such sex difference in  $\gamma_d$  for the BLA hM4Di groups (mean difference in  $\gamma_d = -0.0915$ ; Wilcoxon rank sum test,  $p = 0.24$ ), vIOFC hM4Di groups instead revealed a sex-specific effect between CNO and VEH groups: vIOFC hM4Di females who were administered CNO to inhibit vIOFC had significantly higher values of  $\gamma_d$  compared to those receiving VEH (mean difference in  $\gamma_d = 0.170$ ;  $p = 0.04$ ). This effect was even more pronounced during the first probabilistic R3 (mean difference in  $\gamma_d = 0.125$ ;  $p = 9.42e^{-04}$ ). In contrast, when comparing CNO and VEH in vIOFC hM4Di males, there was no difference in  $\gamma_d$  in either R1 (mean difference in  $\gamma_d = 0.030$ ;  $p = 0.50$ ) or R3 (mean difference in  $\gamma_d = -0.036$ ;  $p = 0.57$ ).

In BLA hM4Di groups, we found a sex-specific effect between CNO and VEH during R3 in the inverse temperature parameter:

BLA hM4Di females administered CNO had significantly higher values of  $\beta$  compared to those receiving VEH (mean difference in  $\beta = 12.5$ ;  $p = 0.02$ ), indicating higher choice consistency or diminished exploration when the better option had reversed, a maladaptive adjustment to reversal. Yet, BLA hM4Di males did not exhibit this difference (mean difference in  $\beta = 7.02$ ;  $p = 0.76$ ). These results based on RL model fitting suggest that the attenuated probabilistic learning in females is mediated by larger  $\beta$  after BLA inhibition and larger  $\gamma_d$  (decreased memory for the unchosen options) after vIOFC inhibition. Importantly, for the vIOFC group, this significant difference emerged due to hM4Di-VEH females exhibiting enhanced memory in R3.

#### Stimulus-based reversal learning

In contrast to the acquisition curves that demonstrated learning of the initial visual discrimination (Fig. 6C), all animals exhibited difficulty with stimulus-based reversal learning, rarely achieving above 60% after 10 sessions (Fig. 7), similar to recent reports (Harris et al., 2021; Ye et al., 2023). Here, due to several nonlinearities, we adhered to the criterion of rats reaching greater than a



**Figure 6.** vIOFC, but not BLA, inhibition impairs adjustment to stimulus-based reversals as measured by probability correct. **A, B,** Trial structure (**A**) and timeline (**B**) of the stimulus-based task. Rats were first surgitized 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), vIOFC hM4Di (middle), and BLA hM4Di (bottom). **D,** Plots of accuracy around reversals were restricted to include only animals that reached greater than 50% running window average for the last 100 trials in initial discrimination. Rats were always tested on a deterministic schedule before a probabilistic one. Shown are subsequent deterministic (100/0) and probabilistic (90/10) reversal “transitions,” three 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 vIOFC abolishes changes in probability correct over the last three (pre-) and first three (post-) reversal sessions, indicating impaired adjustments to reversals. In contrast, BLA inhibition had no clear impact on reversal learning. There was also no effect of CNO in eGFP group learning. **\*\**p* < 0.01** different than eGFP following ANOVA of pre-/post-difference. vIOFC inhibition abolished learning after the first stimulus-based reversal (Fig. 7), but did not significantly impair adjustment around action-based probabilistic reversal as following BLA inhibition (Fig. 8).

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 10 vIOFC hM4Di, 3 of 10 BLA hM4Di, and 5 of 16 eGFP. GLMs were conducted separately in only these “stimulus learners” for accuracy (*probability correct*) by target region comparison but with session instead of trial number as a within-subject factor. Thus, the GLM formula for the full model was as follows:  $\gamma \sim [1 + session \times reversal \text{ number} \times virus \times drug \times drug \text{ order} \times sex + (1 + session \times reversal \text{ number} \times drug | rat)]$ .

Using the above formula, for vIOFC hM4Di comparison with eGFP, we observed several interactions with virus including  $session \times drug \times virus$  ( $\beta_{session \times drug \times virus} = -0.116, p = 1.64e^{-03}$ ),  $session \times drug \times drug \text{ order} \times virus$  ( $\beta_{session \times drug \times drug \text{ order} \times virus} = 0.242, p = 2.47e^{-04}$ ), and  $drug \times virus \times reversal \text{ number}$  ( $\beta_{drug \times virus \times reversal \text{ number}} = -0.133, p = 4.58e^{-03}$ ). With a significant interaction of  $session \times drug \times drug \text{ order} \times reversal \text{ number} \times virus$  ( $\beta_{session \times drug \text{ order} \times drug \times virus \times reversal \text{ number}} = -0.08, p = 4.26e^{-04}$ ), we were justified to analyze reversals separately. We found a significant  $session \times virus$  interaction only in R3 ( $\beta_{session \times virus} = -0.023, p = 0.03$ ). To further probe changes over session, we employed the following formula for each virus separately:  $\gamma \sim [1 + session + (1 + session | rat)]$ . Bonferroni-corrected post hoc comparisons revealed an effect of session in eGFP ( $p < 0.01$ ), but not in hM4Di ( $p = 0.10$ ), indicating that only the eGFP group improved across sessions in R3 (Fig. 7).

For BLA hM4Di compared to eGFP, fewer interactions were observed, but none which included a specific reversal number. We observed an interaction of  $session \times drug \times virus$  ( $\beta_{session \times drug \times virus} = -0.098, p = 0.03$ ); thus, to further probe changes over session, we employed the following formula for each

virus–drug combination separately:  $\gamma \sim [1 + session + (1 + session | rat)]$ . Bonferroni-corrected post hoc comparisons resulted in eGFP-VEH, eGFP-CNO, and hM4Di-VEH exhibiting some improvement across sessions (session, all  $p < 0.01$ ), but not the hM4Di-CNO group (Fig. 7). Thus, BLA inhibition further slowed already incremental learning across all reversals.

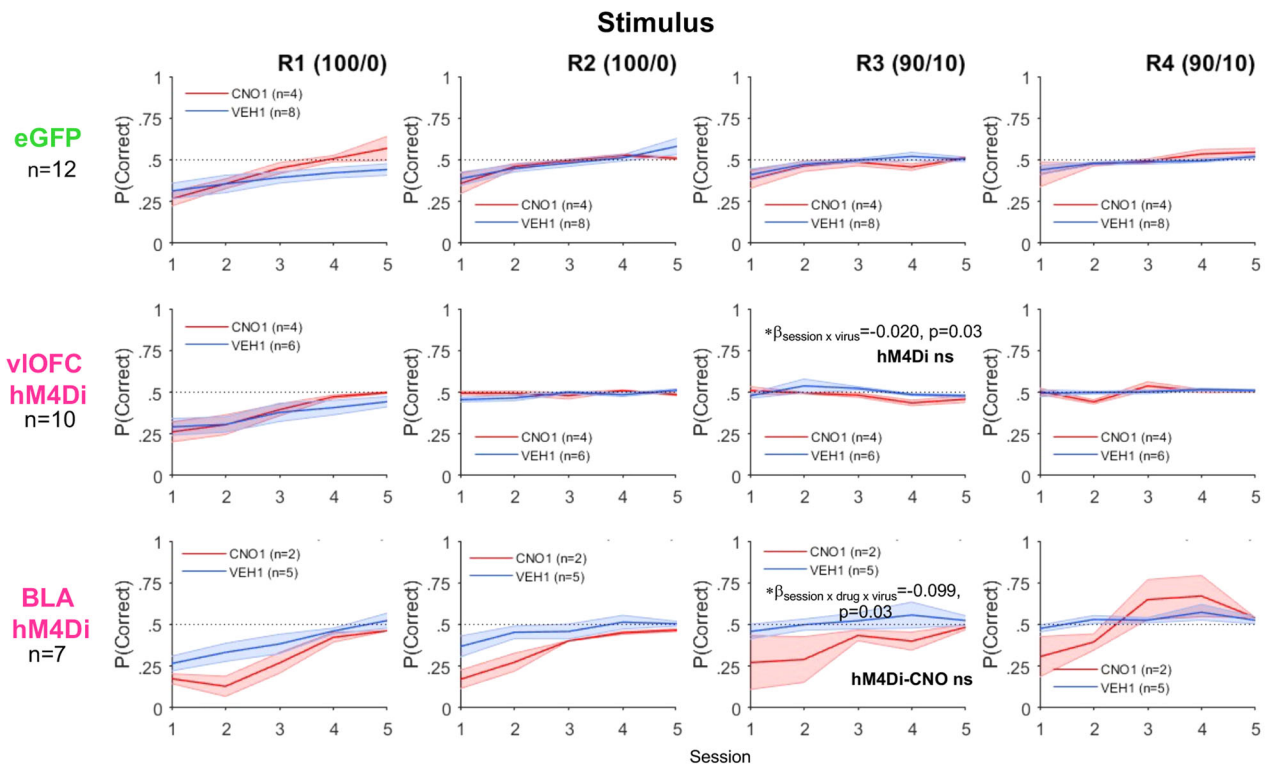
Due to the slow stimulus-based reversal learning, we next assessed performance measured as *probability correct* around reversals (three sessions before and after reversals for stimulus-based learning and one session before and after reversals for action-based learning) to test for adjustments to reversals.

**Accuracy around reversals: probability correct adjustments**  
*Stimulus-based reversal learning*

We analyzed accuracy (*probability correct*) around reversals to assess adjustments to reversals considering that the overall stimulus-based 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. vIOFC hM4Di was significantly different from eGFP for R1-to-R2 [ $F_{(1,18)} = 7.69, p = 0.01$ ] and R2-to-R3 [ $F_{(1,18)} = 7.57, p = 0.01$ ], but not R3-to-R4 [ $F_{(1,18)} = 1.16, p = 0.30$ ; Fig. 6D]. In contrast, BLA hM4Di was not significantly different from eGFP on changes in accuracy around any of the reversals.

*Action-based reversal learning*

For comparison, 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 (Fig. 2), there were no



**Figure 7.** BLA inhibition further slows incremental stimulus-based reversal learning whereas vIOFC inhibition abolishes learning after first reversal. Accuracy in stimulus learners measured by mean probability correct for the first five 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 animals reaching the criterion on initial learning (Fig. 2C), animals exhibited poor stimulus-based reversal learning. Reversal learning was particularly flat for R3 (first probabilistic reversal) following vIOFC inhibition compared to eGFP. Reversal learning across all reversals was especially slow for animals following BLA inhibition. Bonferroni-corrected post hoc comparisons following a mixed-effect GLM wherein a session  $\times$  virus interaction or a session  $\times$  drug  $\times$  virus was found resulting in  $*p < 0.05$  effect of session in eGFP, but not in hM4Di or hM4Di-CNO.

significant effects of virus groups on accuracy changes on any other reversal transition in action-based reversal learning (Fig. 8).

### Strategies around reversals: win-stay, lose-switch

#### 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. vIOFC hM4Di was significantly different from eGFP for *win-stay* R1-to-R2 [ $F_{(1,18)} = 8.97, p < 0.01$ ] and R2-to-R3 [ $F_{(1,18)} = 7.81, p = 0.01$ ] but not R3-to-R4 [ $F_{(1,18)} = 0.40, p = 0.54$ ]. However, though there was a trend for reduced lose-switch in vIOFC hM4Di for R1-to-R2 [ $F_{(1,18)} = 3.64, p = 0.07$ ] and R2-to-R3 [ $F_{(1,18)} = 3.69, p = 0.07$ ], this group did not statistically differ from eGFP on this measure (Fig. 9).

In contrast, BLA hM4Di was significantly different from eGFP on changes in *lose-switch* strategies around only R2-to-R3 [ $F_{(1,15)} = 5.82, p = 0.03$ ]. The results for these adaptive strategies reflect a similar pattern to that observed for *probability correct* for both vIOFC and BLA hM4Di, above.

In summary, our results indicate that BLA, but not vIOFC, is required for learning probabilistic reversals. Conversely, vIOFC, but not BLA, is necessary for rapid adjustments to reversals, generally.

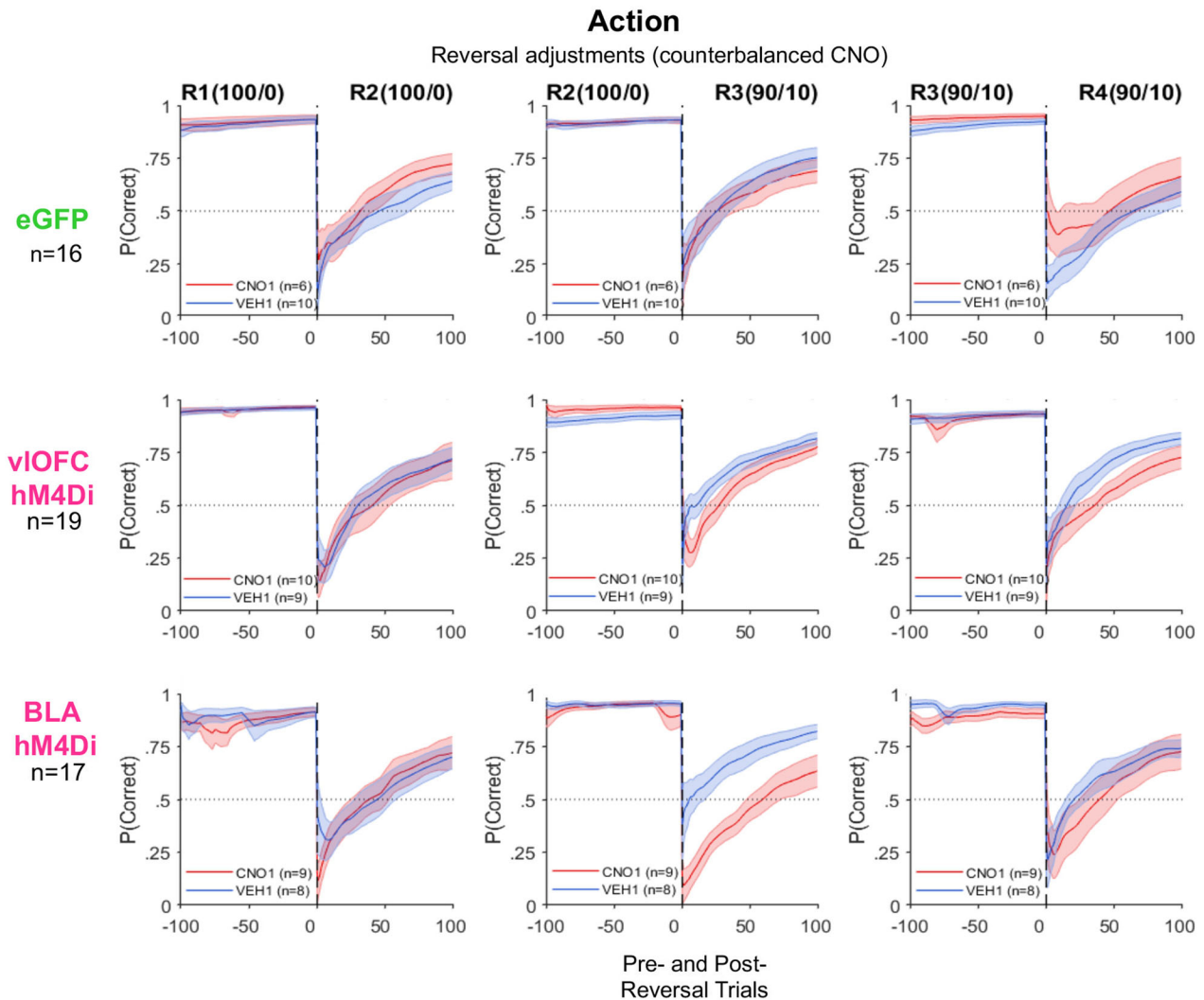
## Discussion

We used a chemogenetic approach to transiently inactivate neurons in either vIOFC 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 targeting projection neurons, we found that rapid adjustments to reversals generally rely on vIOFC, across both stimuli and actions.

In parallel, the specific role of BLA in stimulus- versus 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 prove this, we tested animals on both stimulus- and action-based tasks and found that BLA is required for both stimulus- and action-based reversal learning, with a more pronounced role in probabilistic reversal learning (i.e., detecting meaningful change against a background of uncertainty). This adds to some empirical evidence (Stolyarova and Izquierdo, 2017) and supports its theorized role in learning under unexpected uncertainty (Soltani and Izquierdo, 2019).





**Figure 8.** BLA, but not vOFC, inhibition significantly slowed action-based probabilistic reversal adjustment reflected in accuracy around reversals. Plotted is the probability of choosing the better option [P(correct)] around reversal. 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 the first probabilistic reversal learning as indicated by selectively poor performance on R3 reversal. OFC inhibition resulted in no impairment around these reversals. There was also no effect of CNO in the eGFP group.

As additional motivations for the present study, several reports suggest that neural recruitment in reversal learning may depend on the 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 vOFC to be involved in the rapid adjustment to stimulus-based reversals and in the initial learning of both deterministic and probabilistic learning of actions, whereas BLA was more selectively involved in probabilistic reversal learning, not adjustments to reversals.

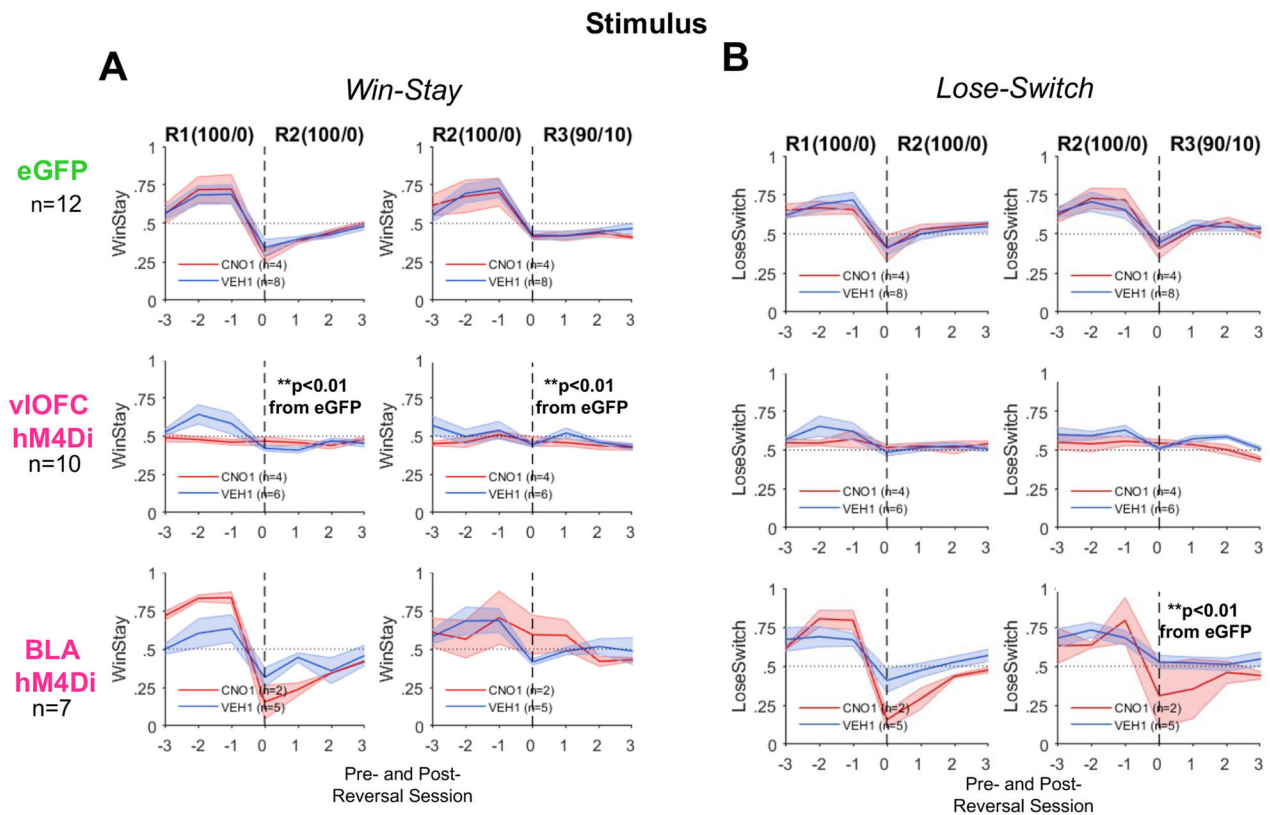
Finally, due to the sparsity of research probing sex differences in flexible learning and decision-making (Orsini and Setlow, 2017; Grissom and Reyes, 2019; Orsini et al., 2022; Cox et al., 2023), where an overwhelming number of reversal learning studies include only males (Schoenbaum et al., 2003; Izquierdo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

action-based reversal learning. We elaborate on these findings within the context of the existing literature below.

#### Preferential recruitment of BLA over vOFC during probabilistic reversal learning

All animals learned to flexibly adjust their responses following deterministic and probabilistic action-based 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 the activation of hM4Di receptors in BLA that was crucial to any impairments observed. vOFC was especially involved in early 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, for example, probabilistic outcomes



**Figure 9.** vIOFC, but not BLA, inhibition impairs adjustments to deterministic stimulus-based reversals as measured by changes in win–stay strategies. Plots of win–stay and lose–switch around reversals were restricted to include only animals that reached greater than 50% running window average for the last 100 trials in initial discrimination. **A**, Inhibition of vIOFC abolishes changes in win–stay over the last three (pre-) and first three (post-) reversal sessions, indicating impaired R1–R2 and R2–R3 transitions. In contrast, BLA inhibition had no impact on performance during these transitions. **B**, Inhibition of BLA attenuates changes in lose–switch over the last three (pre-) and first three (post-) reversal sessions in R2–R3 or the transition from deterministic to probabilistic schedule. There was also no effect of CNO in eGFP group learning. **\*\*** $p < 0.01$  different than eGFP following ANOVA of pre-/post-difference.

(Roesch et al., 2012), and we found evidence of that here. Amygdala-lesioned monkeys are also impaired in action(spatial)-based probabilistic reversal learning, exhibiting a decreased probability of choosing the better option and an 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 R3 suggests it is the reversal superimposed on the misleading feedback that most engages BLA. Here, we also probed if BLA was necessary during the initial learning of probabilistic outcomes, without reversal experience (Fig. 4). Inhibition during initial probabilistic learning did not produce an impairment, indicating BLA is required during the convergence of both probabilistic feedback and shift in contingency (i.e., unexpected uncertainty).

#### vIOFC, but not BLA, is necessary for adjustments to reversals

As described above, unlike the ease of action-based reversal learning, rats exhibited difficulty learning reversals of stimulus–reward contingencies, as previously reported (Harris et al., 2021). Thus, instead of examining acquisition curves that reached asymptotes slightly above chance, we elected to study 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 adjustments

to 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 vIOFC, but not BLA, inhibition produced a failure in detecting the first deterministic and first probabilistic reversal. This pattern was not observed in animals that received VEH during the first deterministic reversal, suggesting vIOFC 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 vIOFC-inhibited animals did not employ effective win–stay strategies after the first reversal. That vIOFC inhibition did not impair the ability to adjust to 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 (Stolyarova and Izquierdo, 2017; Namboodiri et al., 2019; Namboodiri et al., 2021; Jenni et al., 2022), and we add the establishment of expected uncertainty in adjustments to stimulus-based reversals to this evidence.

In contrast, BLA inhibition did not result in any impairment in the ability to adjust to reversals, with the exception of the transition to a probabilistic schedule (i.e., in lose–switch strategies), which lends additional support to the idea that BLA is particularly engaged in adjusting behavior to meaningful changes in the reward environment. Our results based on estimated RL

model parameters suggest differential mechanisms for adjustment to reversals between males and females and following vOFC versus BLA inhibition. Specifically, we found evidence for differential effects on  $\gamma_d$ , where the decay rate for unchosen action values was greater for females than males following vOFC inhibition. This is consistent with a previous study that also reported similar disruption in retention of action values after ablating OFC neurons projecting to BLA (Groman et al., 2019). Notably, Groman et al. (2019) included only male rats and involved a stronger manipulation than our chemogenetic approach (i.e., one that caused pathway-specific neuronal apoptosis). Together with our findings, we can conclude that neurons in OFC, but not BLA, store a memory of action values that are used to adjust to reversals. These results stand in contrast to the maladaptive, increased choice consistency following BLA inhibition that instead reflects poor updating.

### Stimulus-based versus action-based learning

Interestingly, we discovered task order to be significant in rats' ability to learn to discriminate stimuli: stimulus  $\rightarrow$  action was learned much more readily than action  $\rightarrow$  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" versus "where" blocks of trials (Rothenhoefer et al., 2017; Taswell et al., 2021). Nonetheless, learning both types of associations is crucial for the 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 and Koehlin, 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 indicate dissociable roles for vOFC and BLA in flexible learning under uncertainty. BLA is crucial in probabilistic reversal learning or learning of unexpected changes against a background uncertainty, whereas vOFC is important in the establishment of expected uncertainty that enables adjustments to reversals. Interestingly, females exhibited a larger memory decay for the unchosen option following vOFC inhibition than males, indicating a sex-dependent influence on learning under uncertainty. Future investigations could combine targeted causal manipulations with neural correlate approaches to assess the timescales and dynamics of cortico-amygdalar involvement in this learning.

### Code availability

Code and data will be made available upon publication at <https://github.com/izquierdolab> and <https://gin.g-node.org/aizquie>

### References

Alsio J, Lehmann O, McKenzie C, Theobald DE, Searle L, Xia J, Dalley JW, Robbins TW (2021) Serotonergic innervations of the orbitofrontal and medial-prefrontal cortices are differentially involved in visual discrimination and reversal learning in rats. *Cereb Cortex* 31:1090–1105.

Amodeo LR, McMurray MS, Roitman JD (2017) Orbitofrontal cortex reflects changes in response-outcome contingencies during probabilistic reversal learning. *Neuroscience* 345:27–37.

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* 2:100306.

Barreiros IV, Panayi MC, Walton ME (2021a) Organization of afferents along the anterior-posterior and medial-lateral axes of the rat orbitofrontal cortex. *Neuroscience* 460:53–68.

Barreiros IV, Ishii H, Walton ME, Panayi MC (2021b) Defining an orbitofrontal compass: functional and anatomical heterogeneity across anterior-posterior and medial-lateral axes. *Behav Neurosci* 135:165–173.

Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221.

Boulougouris V, Robbins TW (2009) Pre-surgical training ameliorates orbitofrontal-mediated impairments in spatial reversal learning. *Behav Brain Res* 197:469–475.

Boulougouris V, Dalley JW, Robbins TW (2007) Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behav Brain Res* 179:219–228.

Churchwell JC, Morris AM, Heurtelou NM, Kesner RP (2009) Interactions between the prefrontal cortex and amygdala during delay discounting and reversal. *Behav Neurosci* 123:1185–1196.

Corbit LH, Balleine BW (2005) Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci* 25:962–970.

Costa KM, Scholz R, Lloyd K, Moreno-Castilla P, Gardner MPH, Dayan P, Schoenbaum G (2023) The role of the lateral orbitofrontal cortex in creating cognitive maps. *Nat Neurosci* 26:107–115.

Costa VD, Dal Monte O, Lucas DR, Murray EA, Aeverbeck BB (2016) Amygdala and ventral striatum make distinct contributions to reinforcement learning. *Neuron* 92:505–517.

Cox J, et al. (2023) A neural substrate of sex-dependent modulation of motivation. *Nat Neurosci* 26:274–284.

Dalton GL, Wang NY, Phillips AG, Floresco SB (2016) Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. *J Neurosci* 36:1996–2006.

Esber GR, Roesch MR, Bali S, Trageser J, Bissonette GB, Puche AC, Holland PC, Schoenbaum G (2012) Attention-related Pearce-Kaye-Hall signals in basolateral amygdala require the midbrain dopaminergic system. *Biol Psychiatry* 72:1012–1019.

Ghasemi A, Jeddi S, Kashfi K (2021) The laboratory rat: age and body weight matter. *Excli J* 20:1431–1445.

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

Groman SM, Keistler C, Keip AJ, Hammarlund E, DiLeone RJ, Pittenger C, Lee D, Taylor JR (2019) Orbitofrontal circuits control multiple reinforcement-learning processes. *Neuron* 103:734–746 e733.

Harris C, Aguirre C, Kolli S, Das K, Izquierdo A, Soltani A (2021) Unique features of stimulus-based probabilistic reversal learning. *Behav Neurosci* 135:550–570.

Hart EE, Blair GJ, O'Dell TJ, Blair HT, Izquierdo A (2020) Chemogenetic modulation and single-photon calcium imaging in anterior cingulate cortex reveal a mechanism for effort-based decisions. *J Neurosci* 40:5628–5643.

Hervig ME, Fiddian L, Piilgaard L, Bozic T, Blanco-Pozo M, Knudsen C, Olesen SF, Alsio J, Robbins TW (2020) Dissociable and paradoxical roles of rat medial and lateral orbitofrontal cortex in visual serial reversal learning. *Cereb Cortex* 30:1016–1029.

Iordanova MD, Yau JO-Y, McDannald MA, Corbit LH (2021) Neural substrates of appetitive and aversive prediction error. *Neurosci Biobehav Rev* 123:337–351.

Izquierdo A (2017) Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. *J Neurosci* 37:10529–10540.

Izquierdo A, Murray EA (2004) Combined unilateral lesions of the amygdala and orbital prefrontal cortex impair affective processing in rhesus monkeys. *J Neurophysiol* 91:2023–2039.

Izquierdo A, Darling C, Manos N, Pozos H, Kim C, Ostrander S, Cazares V, Stepp H, Rudebeck PH (2013) Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. *J Neurosci* 33:4105–4109.

Janak PH, Tye KM (2015) From circuits to behaviour in the amygdala. *Nature* 517:284–292.



- Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, Averbeck BB (2015) The role of frontal cortical and medial-temporal lobe brain areas in learning a Bayesian prior belief on reversals. *J Neurosci* 35:11751–11760.
- Jenni NL, Rutledge G, Floresco SB (2022) Distinct medial orbitofrontal-striatal circuits support dissociable component processes of risk/reward decision-making. *J Neurosci* 42:2743–2755.
- Jones B, Mishkin M (1972) Limbic lesions and the problem of stimulus-reinforcement associations. *Exp Neurol* 36:362–377.
- Keefner SE, Petrovich GD (2022) Necessity and recruitment of cue-specific neuronal ensembles within the basolateral amygdala during appetitive reversal learning. *Neurobiol Learn Mem* 194:107663.
- Kim J, Ragozzino ME (2005) The involvement of the orbitofrontal cortex in learning under changing task contingencies. *Neurobiol Learn Mem* 83:125–133.
- Lichtenberg NT, Pennington ZT, Holley SM, Greenfield VY, Cepeda C, Levine MS, Wassum KM (2017) Basolateral amygdala to orbitofrontal cortex projections enable cue-triggered reward expectations. *J Neurosci* 37:8374–8384.
- Malvaez M, Shieh C, Murphy MD, Greenfield VY, Wassum KM (2019) Distinct cortical-amygdala projections drive reward value encoding and retrieval. *Nat Neurosci* 22:762–769.
- Namboodiri VMK, Otis JM, van Heeswijk K, Voets ES, Alghorazi RA, Rodriguez-Romaguera J, Mihalas S, Stuber GD (2019) Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. *Nat Neurosci* 22:1110–1121.
- Namboodiri VMK, Hobbs T, Trujillo-Pisanty I, Simon RC, Gray MM, Stuber GD (2021) Relative salience signaling within a thalamo-orbitofrontal circuit governs learning rate. *Curr Biol* 31:5176–5191.e5175.
- Orsini CA, Setlow B (2017) Sex differences in animal models of decision making. *J Neurosci Res* 95:260–269.
- Orsini CA, Truckenbrod LM, Wheeler A-R (2022) Regulation of sex differences in risk-based decision making by gonadal hormones: insights from rodent models. *Behav Process* 200:104663–104663.
- Pachitariu M, Stringer C, Dipoppa M, Schröder S, Rossi LF, Dagleish H, Carandini M, Harris KD (2017) Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv*:061507.
- Piantadosi PT, Lieberman AG, Pickens CL, Bergstrom HC, Holmes A (2018) A novel multichoice touchscreen paradigm for assessing cognitive flexibility in mice. *Learn Mem* 26:24–30.
- Roesch MR, Calu DJ, Esber GR, Schoenbaum G (2010) Neural correlates of variations in event processing during learning in basolateral amygdala. *J Neurosci* 30:2464–2471.
- Roesch MR, Esber GR, Bryden DW, Cerri DH, Haney ZR, Schoenbaum G (2012) Normal aging alters learning and attention-related teaching signals in basolateral amygdala. *J Neurosci* 32:13137–13144.
- Rothenhoefer KM, Costa VD, Bartolo R, Vicario-Feliciano R, Murray EA, Averbeck BB (2017) Effects of ventral striatum lesions on stimulus-based versus action-based reinforcement learning. *J Neurosci* 37:6902–6914.
- Rudebeck PH, Murray EA (2008) Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *J Neurosci* 28:8338–8343.
- Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M (2003) Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learn Mem* 10:129–140.
- Schoenbaum G, Saddoris MP, Stalnaker TA (2007) Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann N Y Acad Sci* 1121:320–35.
- Sias AC, et al. (2021) A bidirectional corticoamygdala circuit for the encoding and retrieval of detailed reward memories. *Elife* 10.
- Soltani A, Izquierdo A (2019) Adaptive learning under expected and unexpected uncertainty. *Nat Rev Neurosci* 20:635–644.
- Soltani A, Koehlin E (2022) Computational models of adaptive behavior and prefrontal cortex. *Neuropsychopharmacology* 47:58–71.
- Stalnaker TA, Roesch MR, Calu DJ, Burke KA, Singh T, Schoenbaum G (2007) Neural correlates of inflexible behavior in the orbitofrontal-amygdala circuit after cocaine exposure. *Ann N Y Acad Sci* 1121:598–609.
- Stolyarova A, Izquierdo A (2017) Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. *Elife* 6.
- Stolyarova A, Rakhshan M, Hart EE, O'Dell TJ, Peters MAK, Lau H, Soltani A, Izquierdo A (2019) Contributions of anterior cingulate cortex and basolateral amygdala to decision confidence and learning under uncertainty. *Nat Commun* 10:4704.
- 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* 39:1720–1728.
- Taswell CA, Costa VD, Basile BM, Pujara MS, Jones B, Manem N, Murray EA, Averbeck BB (2021) Effects of amygdala lesions on object-based versus action-based learning in macaques. *Cereb Cortex* 31:529–546.
- Tye KM, Janak PH (2007) Amygdala neurons differentially encode motivation and reinforcement. *J Neurosci* 27:3937–3945.
- Verharen JPH, den Ouden HEM, Adan RAH, Vanderschuren L (2020) Modulation of value-based decision making behavior by subregions of the rat prefrontal cortex. *Psychopharmacology (Berl)* 237:1267–1280.
- Ward RD, Winiger V, Kandel ER, Balsam PD, Simpson EH (2015) Orbitofrontal cortex mediates the differential impact of signaled-reward probability on discrimination accuracy. *Front Neurosci* 9:230.
- Wassum KM, Izquierdo A (2015) The basolateral amygdala in reward learning and addiction. *Neurosci Biobehav Rev* 57:271–283.
- Winstanley CA, Floresco SB (2016) Deciphering decision making: variation in animal models of effort- and uncertainty-based choice reveals distinct neural circuitries underlying core cognitive processes. *J Neurosci* 36:12069–12079.
- Wright SL, Martin GM, Thorpe CM, Haley K, Skinner DM (2019) Distance and direction, but not light cues, support response reversal learning. *Learn Behav* 47:38–46.
- Ye T, Romero-Sosa JL, Rickard A, Aguirre CG, Wikenheiser AM, Blair HT, Izquierdo A (2023) Theta oscillations in anterior cingulate cortex and orbitofrontal cortex differentially modulate accuracy and speed in flexible reward learning. *Oxford Open Neurosci kvad005*.
- Zimmermann KS, Li CC, Rainnie DG, Ressler KJ, Gourley SL (2018) Memory retention involves the ventrolateral orbitofrontal cortex: comparison with the basolateral amygdala. *Neuropsychopharmacology* 43:373–383.