Dissociable contributions of basolateral amygdala and ventrolateral orbitofrontal cortex to flexible learning under uncertainty

Running Title: *stimulus and action-based reversal learning*

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

[1]Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095
[2]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755
[3]Section of Biosystems and Function, School of Dentistry, University of California, Los Angeles, Los Angeles, CA 90095
[4]Shirley and Stefan Hatos Center for Neuropharmacology, Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA 90095
#co-first authors

Author contributions: CGA and AI designed the research; CGA, JLRS, ZMR, ANT, JM, JP, MG, KD, MG, TY, KE, and JP performed the research; CGA, JHW, JM, JLRS, AS and AI analyzed the data; CGA, JLRS, JHW, PRO, IS, AS, and AI interpreted the data; AI, AS, PRO, and IS acquired funding for the project; CGA and AI wrote the paper; all authors edited the final version.

Correspondence: *Claudia Aguirre, Ph.D. email: cgaguirre@ucla.edu or Alicia Izquierdo, Ph.D. email: aizquie@psych.ucla.edu Ph: +1 310 825 3459

Manuscript Information:
Figures 9

Words: Abstract 248/250
Intro 591/650
Discussion 1513/1500

Abstract

Reversal learning measures the ability to form flexible associations between choice outcomes with stimuli and actions that precede them. This type of learning is thought to rely on several cortical and subcortical areas, including highly interconnected orbitofrontal cortex (OFC) and basolateral amygdala (BLA), and is often impaired in various neuropsychiatric and substance use disorders. However, unique contributions of these regions to stimulus- and action-based reversal learning have not been systematically compared using a chemogenetic approach and particularly before and after the first reversal that introduces new uncertainty. Here, we examined the roles of ventrolateral OFC (vlOFC) and BLA during reversal learning. Male and female rats were prepared with inhibitory DREADDs 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 vlOFC, but not BLA, eliminated adjustments to stimulus-based reversals. Inhibition of BLA, but not vlOFC, selectively impaired action-based probabilistic reversal learning, leaving deterministic reversal learning intact. vlOFC 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 vlOFC in flexible learning and highlight a more crucial role for BLA in learning meaningful changes in the reward environment.

Keywords: **stimulus learning, action learning, deterministic, probabilistic, reward learning, DREADDs**

Significance Statement

Inflexible learning is a feature of several neuropsychiatric disorders. We investigated how the ventrolateral orbitofrontal cortex (vlOFC) and basolateral amygdala (BLA) are involved in learning of stimuli or actions under reinforcement uncertainty. Following chemogenetic inhibition of these regions in both male and females, we measured learning and adjustments to deterministic and probabilistic reversals. For action learning, BLA, but not vlOFC, is needed for probabilistic reversal learning. However, BLA is not necessary for initial probabilistic learning or retention, indicating a critical role for learning of unexpected changes. For stimulus learning, vlOFC, 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.

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 basolateral amygdala (BLA) is an area of interest in reversal learning due 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 vs. actions with the exception of a recent lesion study in rhesus macaques (Taswell et al., 2021). BLA has also not been systematically evaluated for its contributions to deterministic vs. probabilistic schedules, with the exception of another lesion study in monkeys (Costa et al., 2016). The idea that BLA encodes changes in the environment in terms of salience and

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

In parallel, studies with manipulations in rat OFC in reversal learning have included targeting of the entire ventral surface (Izquierdo et al., 2013; Izquierdo, 2017), or more recent systematic comparisons of medial vs. lateral OFC (Hervig et al., 2020; Verharen et al., 2020). Here we examined the role of vlOFC, a subregion not as often probed in reward learning as medial and more (dorso)lateral OFC (cf. Zimmermann et al. (2018)) but also densely-interconnected with BLA (Barreiros et al., 2021a; Barreiros et al., 2021b). 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 vlOFC, but not BLA, inhibition impaired adjustments to deterministic and probabilistic reversals. Conversely, BLA, but not vlOFC, inhibition resulted in significantly slower action-based probabilistic, but not deterministic, reversal learning. Fitting choice data with reinforcement learning models indicated that action-based reversal learning deficits were mediated by a larger memory decay for the unchosen option following vlOFC inhibition, and diminished exploration after reversal following BLA inhibition. These results suggest dissociable roles for BLA in flexible learning under uncertainty, and vlOFC 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) average age post-natal-day (PND) 65 at the start of experiments, with a 280g body weight minimum for males and 240g body weight minimum for females at the time of surgery and the start of the experiment. Rats were approximately PND 100 (emerging adulthood; Ghasemi et al. (2021)) when behavioral testing commenced. Before any treatment, all rats underwent a 3-day acclimation period during which they were pair-housed and given food and water *ad libitum*. During that time, they remained in their home cage with no experimenter interference. Following this 3-day acclimation period, animals were handled for 10 min per animal for 5 consecutive days. During the handling period, the animals were also provided food and water *ad libitum*. After the handling period, animals were individually-housed under standard housing conditions (room temperature 22–24° C) with a standard 12 h light/dark cycle (lights on at 6am). Animals were then surgerized and tested on discrimination and reversal learning 1-week post-surgery. At the point of reversal, they were beyond the 3-week expression time for Designer Receptors Exclusively Activated by Designer Drugs (DREADDs).

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

140

141 **Surgery**

142 *Viral Constructs*

143 Rats were singly-housed and remained in home cages for 4 weeks prior to testing while the
144 inhibitory hM4Di DREADDs expressed in BLA (n=31, 16 females), vlOFC (n=19, 10 females),
145 or eGFP control virus (n=16, 7 females) in these regions. In rats tested on behavior, an adeno-
146 associated virus AAV8 driving the hM4Di-mCherry sequence under the CaMKIIa promoter was
147 used to express DREADDs bilaterally in BLA neurons (0.1 µl, AP = -2.5; ML= ±5; DV = −7.8
148 and 0.2 µl, AP= -2.5; ML= ±5; DV= −8.1, from bregma at a rate of 0.1 µl/min; AAV8-CaMKIIa-
149 hM4D(Gi)-mCherry, Addgene, viral prep #50477-AAV8). In other animals, this same virus
150 (AAV8-CaMKIIa-hM4Di-mCherry, Addgene) was bilaterally infused into two sites in vlOFC (0.2
151 µl, AP = +3.7; ML= ±2.5; DV = −4.6 and 0.15 µl, AP= 4; ML= ±2.5; DV= −4.4, from bregma at
152 a rate of 0.1 µl/min). A virus lacking the hM4Di DREADD gene and only containing the green
153 fluorescent tag eGFP (AAV8-CaMKIIa-eGFP, Addgene) was also infused bilaterally into either
154 BLA (n=7), vlOFC (n=5), or anterior cingulate cortex [(n=5); 0.3 µl, AP = +3.7; ML= ±2.5; DV =
155 −4.6, rate of 0.1 µl/min] as null virus controls. Our vlOFC targeting is most similar to infusion
156 sites reported previously by Dalton et al. (2016) constituting lateral as well as ventral OFC, and
157 0.7 mm more medial than others (Costa et al., 2023). In rats used for *ex vivo* calcium imaging, the
158 same target regions were infused with either GCaMP6f (AAV9-CaMKIIa-GCaMP6f, Addgene), a
159 1:1 combination of GCaMP6f+mCherry (AAV8-CamKIIa-mCherry, Vector BioLabs, #VB1947),
160 or a 1:1 combination of GCaMP6f+hM4Di-mCherry (same as used for behavior, AAV8-
161 CaMKIIa-hM4Di-mCherry, Addgene).

162

163 *Surgical Procedure*

164 Infusions of DREADD or eGFP control virus were performed using aseptic stereotaxic
165 techniques under isoflurane gas (1-5% in O_2) anesthesia prior to any behavioral testing experience.
166 Before surgeries were completed, all animals were administered 5 mg/kg s.c. carprofen (NADA
167 #141–199, Pfizer, Inc., Drug Labeler Code: 000069) and 1cc saline. After being placed in the
168 stereotaxic apparatus (David Kopf; model 306041), the scalp was incised and retracted. The skull
169 was leveled with a +/- 0.3 mm tolerance on the A-P to ensure that bregma and lambda were in the
170 same horizontal plane. Small burr holes were drilled in the skull above the infusion target. Virus
171 was bilaterally infused at a rate of 0.01 µl per minute in target regions (coordinates above). After
172 each infusion, 5 min elapsed before exiting the brain.

173

174 **Histology**

175 At the end of the experiment, rats were euthanized with an overdose of Euthasol (Euthasol,
176 0.8 mL, 390 mg/mL pentobarbital, 50 mg/mL phenytoin; Virbac, Fort Worth, TX), were
177 transcardially perfused, and their brains removed for histological processing. Brains were fixed in
178 10% buffered formalin acetate for 24 h followed by 30% sucrose for 5 days. To visualize hM4Di-
179 mCherry and eGFP expression in BLA or vlOFC cell bodies, free-floating 40-µm coronal sections
180 were mounted onto slides and cover-slipped with mounting medium for DAPI. Slices were
181 visualized using a BZ-X710 microscope (Keyence, Itasca, IL), and analyzed with BZ-X Viewer
182 and analysis software.

183 Reconstructions of viral expressions of hM4Di (magenta) and green fluorescent protein,
184 eGFP (green) across the AP plane (**Fig. 1BE**) were conducted using Photoshop and Illustrator
185 (Adobe, Inc., San Jose, CA). Two independent raters blind to condition then used ImageJ (U. S.

186 National Institutes of Health, Bethesda, Maryland, USA) to trace and quantify pixels at AP +3.7
187 (vlOFC) and AP -2.8 (BLA) for each animal. Three measures were obtained per hemisphere and
188 rater measurements were significantly correlated (Pearson correlation: r = 0.54, p = 4.51e^{-04}).
189 There were no differences in expression level between males and females for pixel count
190 reconstructions [$F_{(1,36)}$ = 2.59, p = 0.12]. Only subjects with bilateral expression were included in
191 behavioral analyses (4 vlOFC and 8 BLA hM4Di rats were excluded due to unilateral expression).
192

193 **Food Restriction**

194 Five days prior to any behavioral testing, rats were placed on food restriction with females
195 on average maintained on 10-12 grams/ day and males given 12-14 grams/ day of chow. Food
196 restriction level remained unchanged throughout behavioral testing, provided animals completed
197 testing sessions. Water remained freely available in the home cage. Animals were weighed every
198 other day and monitored closely to not fall below 85% of their maximum, free-feeding weight.
199

200 **Drug administration**

201 Inhibition of vlOFC or BLA was achieved by systemic administration of clozapine-N-
202 oxide, CNO (i.p., 3mg/kg in 95% saline, 5% DMSO) in animals with DREADDs. Rats with eGFP
203 in these regions underwent identical drug treatment. Rats were randomly assigned to drug
204 treatment groups, irrespective of performance in pretraining. CNO was administered before
205 learning, 30 min prior to behavioral testing. We followed previous work on timing and dose of
206 systemic CNO (Stolyarova, Rakhshan et al., 2019; Hart et al., 2020) and considering the long
207 duration of test sessions. To control for nonspecific effects of injections and handling stress, we
208 also injected animals with saline vehicle (VEH). For reversal learning, to increase power and

209 decrease the number of animals used in experiments, we used a within-subject design for assessing
210 the effects of CNO, with all rats receiving CNO and VEH injections in a counterbalanced order.
211 Thus, for drug administration if a rat received CNO on the first reversal (R1), it was administered
212 VEH on the second reversal (R2), CNO on the third reversal (R3), and VEH on the fourth reversal
213 (R4), or vice versa: VEH on R1, CNO on R2, VEH on R3, and CNO on R4.
214
215 **Behavioral Testing**
216 *Pretraining.* Behavioral testing was conducted in operant conditioning chambers outfitted
217 with an LCD touchscreen opposing the sugar pellet dispenser. All chamber equipment was
218 controlled by customized ABET II TOUCH software (Lafayette Instrument Co., Lafayette, IN).
219 The pretraining protocol, adapted from established procedures (Stolyarova and Izquierdo,
220 2017), consisted of a series of phases: Habituation, Initiation Touch to Center Training (ITCT),
221 Immediate Reward Training (IMT), designed to train rats to nose poke, initiate a trial, and select a
222 stimulus to obtain a reward (i.e., sucrose pellet). Pretraining stages have been reported in detail
223 elsewhere (Stolyarova, Rakhshan et al., 2019). For habituation pretraining, the criterion for
224 advancement was collection of all 5 sucrose pellets. For ITCT, the criterion to the next stage was
225 set to 60 rewards consumed in 45 min. The criterion for IMT was set to 60 rewards consumed in
226 45 min across two consecutive days. After completion of all pretraining schedules, rats were
227 advanced to the discrimination (initial) phase of either the action- or stimulus-based reversal
228 learning task, with the task order counterbalanced (**Fig. 2AB** or **Fig. 6AB**). A subset of animals
229 was tested first on the action-based task (13 vlOFC hM4Di, 10 BLA hM4Di), while others were
230 tested on the stimulus-based task first (9 vlOFC hM4Di, 7 BLA hM4Di, 16 eGFP). Three rats in
231 the vlOFC group completed only the stimulus-based task.

232 *Action-based deterministic discrimination learning.* After completion of either all
233 pretraining schedules or all four reversals of the stimulus-based task, rats were advanced to the
234 discrimination (initial) phase of the action-based task (**Fig. 2A**). Rats were required to initiate a
235 trial by touching the white graphic stimulus in the center screen (displayed for 40 s), and after
236 initiation rats would be presented with two stimuli (i.e., fan or marble) on the left and right side of
237 the screen (displayed for 60 s). Rats could nosepoke either the spatial side rewarded with one
238 sucrose pellet (the better option, $p_R(B)=1.0$;) or the spatial side that went unrewarded (the worse
239 option, $p_R(W)=0.0$). Thus, rats were required to ignore the properties of the stimuli and determine
240 the better rewarded side. If a side was not selected, it was scored as a choice omission, and a 10 s
241 inter-trial interval (ITI) ensued. If a trial was not rewarded, a 5 s time-out would occur, followed
242 by a 10 s ITI. If a trial was rewarded, a 10 s ITI would occur after the reward was collected. The
243 criterion was set to 60 or more rewards consumed and selection of the correct option in 75% of the
244 trials or higher during a 60 min session across two consecutive days. After reaching the criterion
245 for the discrimination phase, the rats were advanced to the reversal phase beginning on the next
246 session. Animals were not administered either CNO or VEH injections during discrimination
247 learning.

248 *Action-based reversal learning.* After the discrimination phase, the rats advanced to the
249 reversal phase. Before a reversal learning session, rats were injected intraperitoneally with either
250 3 mg/kg of CNO or VEH 30 min prior to each reversal testing session. The side previously
251 associated with the $p_R(B)=1.0$ probability was now associated with a $p_R(W)=0.0$ probability of
252 being rewarded, and vice versa. The criterion was the same as the deterministic discrimination
253 phase. After reaching the criterion for the first deterministic reversal phase (i.e., R1), the rats
254 advanced to the second deterministic reversal phase (i.e., R2) beginning on the next session. Rats

255 that had previously received VEH during the first reversal would now receive CNO injections, and
256 vice versa.

257 After completing both deterministic reversal learning phases, rats advanced to the first
258 probabilistic reversal learning phase (i.e., reversal 3, R3). Rats underwent the same injection
259 procedure as the prior reversals. However, the spatial side (i.e., left or right) previously associated
260 with $p_R(B)=1.0$, was now associated with a $p_R(W)=0.1$ probability of being rewarded, whereas the
261 spatial side previously associated with $p_R(W)=0.0$ probability, was now associated with $p_R(B)=0.9$.
262 The criterion was the same as the previous deterministic reversal learning phases. After reaching
263 the criterion for the first probabilistic reversal learning phase (i.e., reversal 3, R3), rats were
264 advanced to the second probabilistic reversal phase (i.e., reversal 4, R4) beginning on the next
265 testing day, where the probabilities would be reversed once again. Rats that had previously
266 received VEH during the first probabilistic reversal now received CNO injections, and vice versa.

267 *Stimulus-based deterministic discrimination learning.* After completion of all pretraining
268 schedules (or all reversals of the action-based task), rats were advanced to the discrimination
269 (initial) phase of learning in which they would initiate a trial by touching a white graphic stimulus
270 in the center screen (displayed for 40 s), and choose between two different visual stimuli
271 pseudorandomly presented on the left and right side of the screen, (**Fig. 6A)**. Stimuli were
272 displayed for 60 s each, randomly assigned as the better or worse stimulus: $p_R(B)=1.0$ or
273 $p_R(W)=0.0$. If a trial was not initiated within 40 s, it was scored as an initiation omission. If a
274 stimulus was not selected, it was scored as a choice omission, and a 10 s ITI ensued. If a trial was
275 not rewarded, a 5 s time-out would occur, followed by a 10 s ITI. Finally, if a trial was rewarded,
276 a 10 s ITI would follow after the reward was collected. A criterion was set to 60 or more rewards
277 consumed and selection of the correct option in 75% of the trials or higher during a 60 min session

278 across two consecutive days. After reaching the criterion for the discrimination phase or if rats
279 were unable to achieve criterion after 10 days, rats were advanced to the reversal phase beginning
280 on the next session. Animals were not administered either CNO or VEH injections during
281 discrimination learning.

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

291 After completing both deterministic reversal learning phases, rats advanced to the first
292 probabilistic reversal learning phase (i.e., reversal 3, R3). The injection procedure remained the
293 same as prior reversals. However, the visual stimulus previously associated with $p_R(B)=1.0$, would
294 now be associated with $p_R(W)=0.1$, whereas the stimulus previously associated with $p_R(W)=0.0$,
295 would now be associated with $p_R(B)=0.9$. The criterion remained the same as prior reversals. After
296 reaching the criterion for the first probabilistic reversal learning phase or if rats were unable to
297 achieve criterion after 10 days, the rats were advanced to the second probabilistic reversal phase
298 (i.e., reversal 4, R4) beginning on the next testing day, where the probabilities would be reversed
299 once again. As above for action-based reversal learning, rats that had previously received VEH
300 during the first probabilistic reversal now received CNO injections, and vice versa.

301 *Action-based probabilistic learning and retention.* We assessed initial probabilistic
302 learning, using the same criterion as above, in a separate group of experimentally-naïve animals
303 expressing hM4Di in BLA (N=14, 7 females; 2 animals were excluded due to unilateral
304 expression). Before a learning session, rats were injected intraperitoneally with either 3 mg/kg of
305 CNO (n=10) or VEH (n=2) 30 min prior to testing. One side of the touchscreen was associated
306 with a better probability of reward ($p_R(B)=0.9$) and the other side, with a worse probability of
307 reward ($p_R(W)=0.1$). Animals were then assessed for their retention of this association on the next
308 session.
309
310 ***Ex vivo* calcium imaging**
311 In N=4 animals (all males), following >3 weeks following stereotaxic viral injections, rats
312 (n=1 rat/brain region/virus combination; n=2-5 slices/rat) were deeply anesthetized with isoflurane
313 (Patterson Veterinary, MA, USA), decapitated and brains submerged in ice cold oxygenated
314 (95/5% O_2/CO_2) slicing artificial cerebrospinal fluid (ACSF) containing (in mM): 62 NaCl, 3.5
315 KCl, 1.25 NaH_2PO_4, 62 choline chloride, 0.5 $CaCl_2$, 3.5 $MgCl_2$, 26 $NaHCO_3$, 5 N-acetyl L-
316 cysteine, and 5 glucose, pH adjusted to 7.3 with KOH. Acutely microdissected vlOFC or BLA
317 slices (300 µM thick) were obtained (VT1200s, Leica, Buffalo Grove, IL) and transferred to room
318 temperature normal ACSF containing (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH_2PO_4, 2 $CaCl_2$, 2
319 $MgCl_2$, 26 $NaHCO_3$, 10 glucose, pH adjusted to 7.3 with KOH, and allowed equilibrate for >1 hr
320 prior to transfer into a perfusion chamber for imaging.
321 Imaging was performed on a Scientifica SliceScope, with imaging components built on an
322 Olympus BX51 upright fluorescence microscope equipped with an sCMOS camera (Hamamatsu
323 Orca Flash 4.0v3). Anatomical regions in brain sections for Ca^{2+} imaging were first identified by
324 brightfield imaging with 780nm LED (Scientifica) illumination. Ca^{2+} imaging was performed

325 using a 40x, 0.80NA water immersion objective (Olympus), continuous 470nm LED illumination
326 (ThorLabs), and a filter cube suitable for GCaMP6f imaging: Excitation: Brightline 466/40,
327 Dichroic: Semrock FF495-Di03, Emission: Brightline 525/50. Slices were housed on poly-D
328 lysine cover slips attached to RC-26G chamber (Warner Instruments, Holliston, MA), which was
329 modified with platinum wires to apply electric field stimulation. Images were acquired continually
330 with 20 ms exposure time. Electric field stimulation was applied at 110 mV (twin pulse every 5s).
331 Temperature of ACSF during the recorded sessions was held at 28° C to minimize bubble
332 formation.
333
334 **Calcium data extraction**
335 Prior to imaging sessions, 40x images of red and green fluorescence were captured and
336 subsequently overlaid for post-hoc genotyping of individual cells (GCaMP6f^{+},
337 GCaMP6f^{+}/hM4Di-mCherry^{+}, or GCaMP6f/mCherry^{+}). Blinded scorers semi-manually curated
338 regions of interest (ROIs) using Python-based Suite2P software (Pachitariu et al., 2017). ROI
339 fluorescence was subtracted from the annular surround fluorescence, low-pass filtered, and
340 transformed to dF/F_0 as previously described (Asrican and Song, 2021) where F_0 is calculated with
341 a boxcar filter with a 200-frame lookback window. dF/F_0 values were clipped between 0 and 9000
342 to eliminate negative changes. Area under the curve and event frequency of each cell was
343 calculated for each drug treatment. A threshold of 0.15 dF/F was used to determine significant
344 events, which is lower than the dF/F of a single *ex vivo* action potential, but significantly above
345 signal to noise in our recorded traces (Tada et al., 2014), **Fig. 1CF**.
346
347 **Data Analyses**

348 MATLAB (MathWorks, Natick, Massachusetts; Version R2021a) was used for all
349 statistical analyses and figure preparation. Data were analyzed with a series of mixed-effects
350 General Linear Models (GLM, *fitglme*) for the discrimination learning phase to establish there
351 were no baseline differences in learning measures between the hM4Di and eGFP animals (i.e.,
352 virus group) prior to any drug treatment for each task separately. Mixed-effects GLMs were also
353 conducted on reversal phases, with all fixed factors included in the model [i.e., reversal number
354 (1-4), virus group (hM4Di, eGFP), drug (CNO, VEH), sex (female, male), drug order (CNO1,
355 VEH1)] and individual rat as a random factor. These GLMs were run for each task type (stimulus-
356 and action-based tasks) separately. Since learning reached asymptote at 5-days for stimulus-based
357 reversal learning, only the first 5 days were included in the GLM. Similarly, since rats typically
358 reached a plateau (and criterion) at 150 trials for action-based reversal learning, we included only
359 the first 150 trials in the GLM. Significant interactions were further analyzed with a narrower set
360 of fixed factors and Bonferroni-corrected post-hoc comparisons. In the instance where sex was
361 found a significant predictor (moderator), sex was entered as a covariate factor in subsequent
362 reversals. Accuracy (probability correct) before and after a reversal (-3 and +3 sessions or -100
363 trials and +100 trials surrounding a reversal) was analyzed using ANOVA with virus group
364 (hM4Di, eGFP) and drug order (CNO1, VEH1) as fixed factors on the average change pre-post
365 reversal. Virus expression level was analyzed with ANOVA by sex (male, female) on pixel counts
366 obtained via imageJ. Pearson correlation coefficients (*corrcoef*) were analyzed for inter-rater
367 reliability of viral expression quantification. Wilcoxon rank sum tests were used to compare
368 reinforcement learning parameters between groups.

369 Dependent measures for learning included probability of choosing the correct or better
370 option, probability of win-stay, and probability of lose-switch. Probability of win-stay and lose-

switch adaptive strategies were calculated for the stimulus-based task such that each trial was classified as a *win* if an animal received a sucrose pellet, and as a *loss* if no reward was delivered. Statistical significance was noted when p-values were less than 0.05. All Bonferroni post-hoc tests were corrected for number of comparisons.

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

Reinforcement Learning Models

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

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

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

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

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

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

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

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

We used the standard maximum likelihood estimation method to fit choice data and estimate the parameters for each session of the experiment. The values of the learning rate α and decay rate γ_d were bounded between 0 and 1, and β was bounded between 1 and 100. Initial parameter values were selected from this range, and fitting was performed using the MATLAB function *fmincon*. For each set of parameters fitted to each session, we repeated 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_{decay}), both choice options were considered unchosen for those trials and both of the value estimates decayed passively according to Equation (2).

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

416 $$AIC = -2 * LL + 2 * k, \quad (4)$$

417 $$BIC = -2 * LL + \ln(n) * k, \quad (5)$$

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

420

421 **Results**

422 ***Ex vivo calcium imaging in slices***

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

430 In vlOFC, there was also no significant effect of CNO (10μM) on Ca^{2+} events for neurons
431 expressing GCaMP6f or GCaMP6f+mCherry (**Fig. 1F**). However, in $CaMKII^{+}$ vlOFC neurons
432 expressing GCaMP6f+hM4Di there was a decrease in the frequency of Ca^{2+} events during CNO
433 application. A 2-way ANOVA revealed a significant drug x virus interaction [$F_{(2,400)} = 8.349$, $p<$
434 0.001], with multiple comparisons test resulting in decreased Ca^{2+} events in GCaMP6f+hM4Di
435 following CNO ($p = 0.02$), and increased activity in GCaMP6f expressing neurons after CNO (p
436 $= 0.02$).

437

438 ***Discrimination learning: eGFP controls***

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

447

448 ***Discrimination learning: hM4Di vs. eGFP***

449 For the action-based task, there were no significant effects of virus or virus interactions for
450 vlOFC vs. eGFP on *probability correct* (β_{virus} = -0.002, p = 0.96), with similar findings for the
451 comparison of BLA vs. eGFP (β_{virus} = -0.02, p= 0.85; **Fig. 2C**). All animals met criterion very
452 quickly (~2 days), thus, we compared trials to reach 75% criterion (i.e., probability of choosing
453 the correct side). Both hM4Di virus groups performed comparably [M±SEM: vlOFC hM4Di
454 (69.6±15.4, BLA hM4Di (80.9±21.5)], whereas the eGFP group met criterion within fewer trials
455 (59.5±18.6), but the difference was not statistically significant [vlOFC hM4Di vs. eGFP: β_{virus} =
456 33.2, p = 0.34; BLA hM4Di vs. eGFP: β_{virus} = 51.3, p = 0.21].

457 For the stimulus-based task, there were also no significant effects of virus or virus
458 interactions for either vlOFC vs. eGFP on *probability correct* (β_{virus} = -0.09, p = 0.10), or for BLA
459 vs. eGFP (β_{virus} = -0.06, p = 0.20; **Fig. 6C**). The animals on average took approximately ~6 days

460 to meet criterion regardless of virus group [M±SEM: vlOFC hM4Di (6.1±0.7), BLA hM4Di
461 (6.5±1.2), eGFP (6.9±0.7)].

462 Given the poorer learning in the stimulus-based task, we evaluated whether this was due to
463 the order of task administered [i.e., Stimulus → Action or Action → Stimulus]. To test whether
464 learning was influenced by task order, we analyzed *probability correct* during initial
465 discrimination learning for the stimulus-based task, which resulted in no effect of task order (β_{order}
466 $= 0.03$, $p = 0.33$), but a significant task order x session interaction ($\beta_{order\ x\ session} = -0.03$, $p = 0.002$).
467 Thus, subsequent analyses were conducted with task order analyzed separately by session, which
468 revealed that animals administered the Action → Stimulus task order exhibited poorer learning
469 across sessions ($\beta_{session} = 0.01$, $p = 0.05$), compared to those administered the Stimulus→ Action
470 task order ($\beta_{session} = 0.04$, $p < 0.0001$). To further address the possibility that any effects on reversal
471 learning could be due to task order instead of DREADDs inhibition, we conducted a separate
472 assessment of learning in animals administered the Action-based task first (see below).

473

474 ***Accuracy across trials and sessions: Reversal learning***

475 *Action-based reversal learning.* Mixed-effects GLMs were used to analyze *probability*
476 *correct,* our primary measure of accuracy, with drug order, virus, and sex as between-subject
477 factors, trial number, reversal number, and drug as within-subject factors, and individual rat as
478 random factor. GLMs were conducted separately by target region (BLA vs. eGFP and vlOFC vs.
479 eGFP), using the following formula for the full model: $\gamma \sim$ [*1 + trial number + reversal number **
480 *virus *drug * drug order * sex + (1 + trial number + reversal number * drug| rat)*].

481 Starting with the full model above for the comparison of BLA with eGFP, we found that
482 trial number was a significant predictor of *probability correct* ($\beta_{trial\ number} = 0.003$, $p = 3.26e^{-81}$). We

483 also observed a sex x drug order ($\beta_{sex\ x\ drug\ order}$ =0.63, p = 0.01) and sex x virus x drug order ($\beta_{sex\ x}$
484 $_{virus\ x\ drug\ order}$ =-0.72, p = 0.046) interaction. To further probe "first reversals" we included only R1
485 and R3 in the above model. Accordingly, drug order did not vary across R1 and R3. In the analysis
486 of first experiences with deterministic R1 and probabilistic R3, trial number was also a significant
487 predictor of accuracy ($\beta_{trial\ number}$ =0.003, p = 2.59e^{-49}), along with a virus x drug x reversal number
488 interaction ($\beta_{virus\ x\ drug\ x\ reversal\ number}$ =-0.19, p = 0.04), which justified further analysis of R1 and R3,
489 separately. Follow-up GLMs were conducted for R1 and R3 *probability correct*, with the following
490 formula: $\gamma \sim [1 + virus * drug + (1 \mid rat)]$. A significant virus x drug interaction was obtained only
491 for R3: ($\beta_{virus\ x\ drug}$ =-0.22, p = 0.03), **Fig. 2D**. Bonferroni-corrected post-hoc comparisons revealed
492 an effect of CNO in hM4Di (p = 2.33e^{-04}), not in eGFP (p = 1.0), with both males and females
493 exhibiting impaired learning of probabilistic R3 following BLA inhibition (**Fig. 3B**). Importantly,
494 stimulus-naïve animals also demonstrated impaired probabilistic reversal learning (β_{drug} =-0.23, p
495 = 0.02) (**Fig. 4A**), indicating this slower learning was not due to previous training history.

496 For the comparison of vlOFC with eGFP, trial number was a significant predictor of
497 *probability correct* ($\beta_{trial\ number}$ =0.003, p = 6.39e^{-75}). We also observed a sex x drug order ($\beta_{sex\ x\ drug}$
498 $_{order}$ =0.60, p = 0.03) interaction (**Fig. 2D**). Follow-up GLMs were conducted for accuracy in
499 females and males separately, with the following formula: $\gamma \sim [1 + drug\ order + (1 \mid rat)]$. A
500 significant drug order effect (CNO first > VEH first) was obtained for females: ($\beta_{drug\ order}$ =-0.07,
501 p = 0.01), but not males (**Fig. 3**). Due to finding no interactions of reversal number with any other
502 predictor, we were not justified to look further at individual reversals. Altogether, there was no
503 effect of vlOFC inhibition on action-based reversal learning: all animals eventually reached
504 asymptote and criterion over a comparable number of trials.

505 Since OFC has been implicated in early reversal learning both classically and recently
506 (Jones & Mishkin, 1972; Schoenbaum et al. 2007; Amodeo et al. 2017), we conducted additional
507 analyses using the same full-model formula $\gamma \sim [1 + trial\ number + reversal\ number * virus * drug$
508 $* drug\ order * sex + (1 + trial\ number + reversal\ number * drug | rat)]$, but this time restricted to
509 early trials (first 10) of each reversal. Within this early phase, there was a significant sex x virus x
510 drug order ($\beta_{sex\ x\ virus\ x\ drug\ order}$=-0.94, p = 0.03) and sex x virus x drug x drug order interaction (β_{sex}
511 $_{x\ virus\ x\ drug\ x\ drug\ order}$=1.05, p = 0.04). When sex was entered as a covariate, there was a marginally
512 significant interaction of virus x drug order for all reversals ($\beta_{virus\ x\ drug\ order}$=-0.22, p = 0.056), with
513 a sex difference observed only in the hM4Di group (p = 0.03), not in eGFP (p = 0.44). Thus, early
514 reversal adjustment in female rats was more adversely affected by previous OFC inhibition (i.e.,
515 if they received CNO first), than in males (**Fig. 3**).

516 *Action-based probabilistic learning and retention.* To further probe if the effect of BLA
517 inhibition on probabilistic reversal learning was specific to reversal, and not initial learning, we
518 administered CNO or VEH before initial 90/10 learning in a separate group of experimentally
519 naïve animals transfected with hM4Di in BLA. To analyze the effect of drug on *probability correct*
520 we used the formula: $\gamma \sim [1 + trial\ number + drug + (1 + trial\ number | rat)]$. Only an effect of
521 trial number was found ($\beta_{trial\ number}$=0.001, p = 1.7e^{-04}). There was no significant effect of drug on
522 learning in either session 1 (initial learning) or session 2 (retention), with or without sex entered
523 in the model as a covariate. Thus, all rats demonstrated learning and full retention on the next day
524 (**Fig. 4B**).

525 *Fitting choice behavior in R1 and R3 with Reinforcement learning models.* To gain more
526 insight into the effect of vlOFC and BLA inhibition on deterministic reversal learning (R1) and
527 probabilistic reversal learning (R3) and their potential underlying mechanisms, we next compared

528 the estimated model parameters from reinforcement learning models (RL, RL_{decay}). Comparing the
529 goodness of fit between the two models, we found that the second model with the decay parameter
530 (RL_{decay}) better accounted for the animals' choice behavior as indicated by significantly lower AIC
531 (paired sample t-test; $t(1334) = 4.613$, $p = 4.34e^{-06}$). In contrast, the overall mean BIC value was
532 significantly lower for the first model ($t(1334) = -2.623$, $p = 8.00\ e^{-03}$). We focused on the estimated
533 parameters from the $RL1_{decay}$ model only (**Fig. 5**), as AIC better distinguished the better-fitting
534 model and allowed estimation of the additional γ_d parameter.

535 Comparison of the estimated parameters across groups revealed that female and male rats
536 differed mainly in the decay parameter γ_d, which governs the amount of passive decay or forgetting
537 in the value estimate of the unchosen option. During the first deterministic reversal (R1, 100/0),
538 eGFP females showed overall significantly lower values of γ_d than eGFP males (mean difference
539 in γ_d = -0.0982; Wilcoxon rank sum test, $p = 3.40e^{-06}$), suggesting a different mechanism of
540 adjustment to the reversal between females and males. While there was no clear evidence for such
541 sex difference in γ_d for the BLA hM4Di groups (mean difference in γ_d = -0.0915; Wilcoxon rank
542 sum test, $p = 0.24$), vlOFC hM4Di groups instead revealed a sex-specific effect between CNO and
543 VEH groups: vlOFC hM4Di females who were administered CNO to inhibit vlOFC had
544 significantly higher values of γ_d compared to those receiving VEH (mean difference in $\gamma_d = 0.170$;
545 $p = 0.04$). This effect was even more pronounced during the first probabilistic reversal R3 (mean
546 difference in $\gamma_d = 0.125$; $p = 9.42e^{-04}$). In contrast, when comparing CNO and VEH in vlOFC
547 hM4Di males there was no difference in γ_d in either R1 (mean difference in $\gamma_d = 0.030$; $p = 0.50$)
548 or R3 (mean difference in γ_d = -0.036; $p = 0.57$).

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

551 higher values of β compared to those receiving VEH (mean difference in β = 12.5; p = 0.02),
552 indicating higher choice consistency or diminished exploration when the better option had
553 reversed, a maladaptive adjustment to reversal. Yet, BLA hM4Di males did not exhibit this
554 difference (mean difference in β = 7.02; p = 0.76). These results based on RL model fitting suggest
555 that the attenuated probabilistic learning in females is mediated by larger β after BLA inhibition
556 and larger γ_d (decreased memory for the unchosen options) after vlOFC inhibition. Importantly,
557 for vlOFC group, this significant difference emerged due to hM4Di VEH females exhibiting
558 enhanced memory in R3.

559 *Stimulus-based reversal learning.* In contrast to the acquisition curves that demonstrated
560 learning of the initial visual discrimination (**Fig. 6C**), all animals exhibited difficulty with
561 stimulus-based reversal learning, rarely achieving above 60% after 10 sessions (**Fig. 7**), similar to
562 recent reports (Harris, Aguirre et al., 2021; Ye et al., 2023). Here, due to several non-learners, we
563 adhered to the criterion of rats reaching greater than a 50% running window average for the last
564 100 trials in discrimination, for inclusion in subsequent reversal learning analyses. The following
565 numbers did not meet this criterion and were excluded from these groups: 0 of 10 vlOFC hM4Di,
566 3 of 10 BLA hM4Di, and 5 of 16 eGFP. GLMs were conducted separately in only these 'stimulus
567 learners' for accuracy (*probability correct*) by target region comparison, but with session instead
568 of trial number as a within-subject factor. Thus, the GLM formula for the full model was as
569 follows: $\gamma \sim [1 + session * reversal\ number * virus * drug * drug\ order * sex + (1 + session *$
570 $reversal\ number * drug | rat)]$.

571 Using the above formula, for vlOFC hM4Di comparison with eGFP, we observed several
572 interactions with virus including: session x drug x virus ($\beta_{session\ x\ drug\ x\ virus}$ = -0.116, p = 1.64e^{-03}),
573 session x drug x drug order x virus ($\beta_{session\ x\ drug\ x\ drug\ order\ x\ virus}$ = 0.242, p = 2.47e^{-04}) and drug x virus

574 x reversal number ($\beta_{\text{drug x virus x reversal number}}$ = -0.133, p = 4.58e^{-03}). With a significant interaction of
575 session x drug x drug order x reversal number x virus ($\beta_{\text{session x drug order x drug x virus x reversal number}}$ = -
576 0.08, p = 4.26e^{-04}), we were justified to analyze reversals separately. We found a significant session
577 x virus interaction only in R3 ($\beta_{\text{session x virus}}$ = -0.023, p = 0.03). To further probe changes over
578 session we employed the following formula for each virus separately: $\gamma \sim [1 + session + (1 +$
579 $session|\ rat)]$. Bonferroni-corrected post-hoc comparisons revealed an effect of session in eGFP
580 ($p < 0.01$), but not in hM4Di (p = 0.10), indicating that only the eGFP group improved across
581 session in R3 (**Fig. 7**).

582 For BLA hM4Di compared to eGFP, fewer interactions were observed, but none which
583 included specific reversal number. We observed an interaction of session x drug x virus ($\beta_{\text{session x}}$
584 $_{\text{drug x virus}}$ = -0.098, p = 0.03), thus to further probe changes over session, we employed the following
585 formula for each virus-drug combination separately: $\gamma \sim [1 + session + (1 + session|\ rat)]$.
586 Bonferroni-corrected post-hoc comparisons resulted in eGFP-VEH, eGFP-CNO, and hM4Di-VEH
587 exhibiting some improvement across sessions (session, all $p < 0.01$), but not the hM4Di-CNO
588 group (**Fig. 7**). Thus, BLA inhibition further slowed already incremental learning across all
589 reversals.

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

594

595 ***Accuracy around reversals: Probability correct adjustments***

596 *Stimulus-based reversal learning.* We analyzed accuracy (*probability correct*) around
597 reversals to assess adjustments to reversals considering that the overall stimulus-based learning
598 was modest. ANOVAs with virus and drug order as between subject factors were conducted on
599 the mean change in accuracy between one reversal and the next. vlOFC hM4Di was significantly
600 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],
601 but not R3-to-R4 [$F_{(1,18)}$= 1.16, p = 0.30], **Fig. 6D**. In contrast, BLA hM4Di was not significantly
602 different from eGFP on changes in accuracy around any of the reversals.
603
604 *Action-based reversal learning.* For comparison, we also assessed accuracy (*probability*
605 *correct*) around reversals for action-based reversal learning. As above, ANOVAs with virus and
606 drug order as between subject factors were conducted on the mean accuracy change between one
607 reversal and the next. Other than confirming the probabilistic reversal learning (R3) impairment
608 for BLA hM4Di (**Fig. 2**), there were no significant effects of virus groups on accuracy changes on
609 any other reversal transition in action-based reversal learning (**Fig. 8).**
610
611 ***Strategies around reversals: Win-Stay, Lose-Switch***
612 *Stimulus-based reversal learning.* We also analyzed adaptive response strategies (*Win-Stay*
613 and *Lose-Switch*) around reversals. ANOVAs with virus and drug order as between subject factors
614 were conducted on mean Win-Stay or Lose-Switch between one reversal and the next. vlOFC
615 hM4Di was significantly different from eGFP for *Win-Stay* R1-to-R2 [$F_{(1,18)}$ =8.97, p < 0.01] and
616 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
617 there was a trend for reduced Lose-Switch in vlOFC hM4Di for R1-to-R2 [$F_{(1,18)}$= 3.64, p = 0.07]

618 and R2-to-R3 [$F_{(1,18)}$= 3.69, p = 0.07], this group did not statistically differ from eGFP on this
619 measure, **Fig. 9**.
620 In contrast, BLA hM4Di was significantly different from eGFP on changes in *Lose-Switch*
621 strategies around only R2-to-R3 [$F_{(1,15)}$= 5.82, p = 0.03]. The results for these adaptive strategies
622 reflect a similar pattern to that observed for *probability correct* for both vlOFC and BLA hM4Di,
623 above.
624 In summary, our results indicate that BLA, but not vlOFC, is required for learning
625 probabilistic reversals. Conversely, vlOFC, but not BLA, is necessary for rapid adjustments to
626 reversals, generally.
627
628
629 **Discussion**
630
631 We used a chemogenetic approach to transiently inactivate neurons in either vlOFC or
632 BLA to assess how these regions are involved in different aspects of reversal learning. Although
633 the role of OFC in reversal learning has been instantiated in different paradigms using visual
634 stimuli and cues (Izquierdo et al., 2013; Piantadosi et al., 2018; Hervig et al., 2020; Alsio et al.,
635 2021) as well as olfactory ones (Schoenbaum et al., 2003; Kim and Ragozzino, 2005), several
636 groups also report a strong role for OFC in action (spatial)-based reversal learning (Dalton et al.,
637 2016; Groman et al., 2019; Verharen et al., 2020). Almost all of these reversal learning
638 investigations have involved irreversible lesions or baclofen/muscimol inactivations of OFC.
639 Testing both types with a chemogenetic approach targeting projection neurons, here we found that
640 rapid adjustments to reversals generally rely on vlOFC, across both stimuli and actions.
641 In parallel, the specific role of BLA in stimulus- vs. action-based reversal learning is poorly
642 understood given mixed results (Schoenbaum et al., 2003; Izquierdo and Murray, 2004;

643 Churchwell et al., 2009; Hervig et al., 2020). Recent studies suggest amygdala may be involved in
644 both types of learning (Taswell et al., 2021; Keefer and Petrovich, 2022) as BLA activity is
645 modulated by violations in reward expectations generally, which are not association-specific
646 (Roesch et al., 2012). To probe this, we tested animals on both stimulus- and action-based tasks
647 and found that BLA is required for both stimulus- and action-based reversal learning, with a more
648 pronounced role in probabilistic reversal learning (i.e., detecting meaningful change against a
649 background of uncertainty). This adds to some empirical evidence (Stolyarova & Izquierdo, 2017)
650 and supports its theorized role in learning under unexpected uncertainty (Soltani & Izquierdo,
651 2019).

652 As additional motivations for the present study, several reports suggest that neural
653 recruitment in reversal learning may depend on certainty of rewards (Boulougouris et al., 2007;
654 Boulougouris and Robbins, 2009; Ward et al., 2015; Costa et al., 2016; Dalton et al., 2016;
655 Piantadosi et al., 2018; Verharen et al., 2020). To further understand this, we tested animals on
656 both deterministic (100/0) and probabilistic reversals (90/10). We found vlOFC to be involved in
657 the rapid adjustment to stimulus-based reversals and in the initial learning of both deterministic
658 and probabilistic learning of actions, whereas BLA was more selectively involved in probabilistic
659 reversal learning, not adjustments to reversals.

660 Finally, due to the sparsity of research probing sex differences in flexible learning and
661 decision making (Orsini and Setlow, 2017; Grissom and Reyes, 2019; Orsini et al., 2022; Cox et
662 al., 2023), where an overwhelming number of reversal learning studies include only males
663 (Schoenbaum et al., 2003; Izquierdo et al., 2013; Dalton et al., 2016; Groman et al., 2019; Hervig
664 et al., 2020; Verharen et al., 2020), we included both male and female rats here. We found sex-

665 dependent contributions of vlOFC in early action-based reversal learning. We elaborate on these
666 findings within the context of the existing literature below.
667
668 ***Preferential recruitment of BLA over vlOFC during probabilistic reversal learning***
669 All animals learned to flexibly adjust their responses following deterministic and
670 probabilistic action-based reversals, indicating successful remapping of reward contingencies as
671 accuracy increased across trials. Importantly, we found no effect of CNO in eGFP animals,
672 suggesting that it was activation of hM4Di receptors in BLA that were crucial to any impairments
673 observed. vlOFC was especially involved in early learning of first deterministic (R1) and
674 probabilistic reversal (R3). This is consistent with findings following pharmacological
675 inactivations or lesions of OFC (Boulougouris et al., 2007; Boulougouris and Robbins, 2009;
676 Dalton et al., 2016; Piantadosi et al., 2018; Verharen et al., 2020).
677 BLA inhibition was not expected to impair deterministic reversal learning as it is thought
678 to be mostly recruited when there is some level of uncertainty, e.g., probabilistic outcomes (Roesch
679 et al., 2012), and we found evidence of that here. Amygdala-lesioned monkeys are also impaired
680 on action(spatial)-based probabilistic reversal learning, exhibiting decreased probability of
681 choosing the better option, and increased switching behavior following negative outcomes
682 (Taswell et al., 2021). BLA may indeed be critical in generating prediction error signals following
683 changes in reward associations (Esber et al., 2012; Roesch et al., 2012; Iordanova et al., 2021),
684 with particular involvement in detecting unexpected upshifts or downshifts in value (Roesch et al.,
685 2010; Stolyarova and Izquierdo, 2017). Our finding of attenuated learning of probabilistic reversal
686 R3 suggests it is the reversal superimposed on the misleading feedback that most engages BLA.
687 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).

vlOFC, 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, Aguirre et al., 2021). Thus, instead of examining acquisition curves which reached asymptote 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 vlOFC, but not BLA, inhibition produced a failure in detecting first deterministic and first probabilistic reversal. This pattern was not observed in animals that received VEH during the first deterministic reversal, suggesting vlOFC needs to be "online" when experiencing a reversal for the very first time as this determines how flexibly animals respond to future reversals. Employment of adaptive strategies matched this effect, such that vlOFC-inhibited animals did not employ effective win-stay strategies after the first reversal. That vlOFC 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 & Izquierdo, 2017; Namboodiri et al., 2019;

711 Namboodiri et al., 2021; Jenni et al., 2022), and we add the establishment of expected uncertainty
712 in adjustments to stimulus-based reversals to this evidence.

713 In contrast, BLA inhibition did not result in any impairment in the ability to adjust to
714 reversals, with the exception of the transition to a probabilistic schedule (i.e., in lose-switch
715 strategies), which lends additional support to the idea that BLA is particularly engaged in adjusting
716 behavior to meaningful changes in the reward environment. Our results based on estimated RL
717 model parameters suggests differential mechanisms for adjustment to reversals between males and
718 females and following vlOFC vs. BLA inhibition. Specifically, we found evidence for differential
719 effects on γ_d, where the decay rate for unchosen action values was greater for females than males
720 following vlOFC inhibition. This is consistent with a previous study which also reported similar
721 disruption in retention of action values after ablating OFC neurons projecting to BLA (Groman et
722 al., 2019). Notably, Groman et al. (2019) included only male rats and involved a stronger
723 manipulation than our chemogenetic approach (i.e., one that caused pathway-specific neuronal
724 apoptosis). Together with our findings, we can conclude that neurons in OFC, but not BLA, store
725 a memory of action values that are used to adjust to reversals. These results stand in contrast to the
726 maladaptive, increased choice consistency following BLA inhibition that instead reflects poor
727 updating.

728

729 ***Stimulus-based vs. action-based learning***

730 Interestingly, we discovered task order to be significant in rats' ability to learn to
731 discriminate stimuli: stimulus→action was learned much more readily than action→stimulus. This
732 can be explained by noting that rats are heavily biased to acquire spatial associations (Wright et
733 al., 2019), and reinforcing this already-strong learning likely inhibits the ability to learn

associations where spatial information should be ignored. In contrast, nonhuman primates are able to quickly transition between “what” vs. “where” blocks of trials (Rothenhoefer et al., 2017; Taswell et al., 2021). Nonetheless, learning both types of associations are crucial for flexibility required in naturalistic environments and thus, it is important to examine how stimulus-based and action-based learning systems interact with each other (Soltani & Koechlin, 2022). Moreover, although the role of OFC in stimulus- or cue-based reversal learning has been probed using olfactory and visual stimuli, more viral-mediated approaches employing targeted chemogenetic and optogenetic manipulations across sensory modalities in both males and females are warranted.

Conclusion

The present results indicate dissociable roles for vlOFC and BLA in flexible learning under uncertainty. BLA is crucial in probabilistic reversal learning, or learning of unexpected changes against a background uncertainty, whereas vlOFC 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 vlOFC 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.

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

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

The authors declare no competing financial interests.

768 **Figure Captions**

769

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

788

789 **Figure 2. BLA inhibition impairs action-based probabilistic reversal learning, whereas vlOFC**
790 **inhibition only slows early adjustment to reversals. (A-B)** Trial structure (A) and timeline (B) of the
791 action-based task. Rats were first surgerized with either hM4Di DREADDs on a CaMKII promoter or eGFP
792 null virus, also on the same promoter. Rats were allowed to recover for 1 week before testing on a stimulus-
793 or action-based reversal learning task. **(C)** Initial learning of a rewarded side. **(D)** Learning during
794 subsequent deterministic (100/0) and probabilistic (90/10) reversal**.** Plots show cumulative P(Correct) for
795 first 150 trials with a sliding window of 10 trials. Drug order was counterbalanced such that on R2 and R4
796 animals received VEH if they were administered CNO first on R1 and R3, and vice versa. There was no
797 effect of inhibition on learning in R2 and R4 (not shown). There was no effect of CNO on learning in the
798 eGFP group. *$p<0.05$ significant sex by drug order interaction (see **Fig. 3** for learning curves plotted by
799 sex). Bonferroni-corrected post-hocs following mixed-effects GLM with sex as a covariate fixed factor
800 wherein a drug x virus interaction was found resulted 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 vlOFC inhibition than in males, whereas BLA role in probabilistic reversal learning was not sex-dependent. Plotted is the accuracy for 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 vlOFC lowered P(Correct) in early first deterministic R1 and first probabilistic reversal R3 in females but not males, whereas BLA inhibition attenuated probabilistic reversal learning in both males and females. Bonferroni-corrected post-hocs upon GLM result of sex x drug order interaction resulted in an effect of drug order only in vlOFC 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. **$p=0.01$, *$p<0.05$. Reinforcement learning model fit to choice behavior also indicated that chemogenetic inhibition of vlOFC increased the decay rate (decreased memory) of the unchosen option (γ_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.

Figure 5. Reinforcement learning 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. α = learning rate, β = inverse temperature, γd= decay rate for unchosen option. Drug order was counterbalanced such that on R2 and R4 (not shown) animals received VEH if they were administered CNO first on R1 and R3, and vice versa. Female eGFP group exhibited overall lower γd than male eGFP rats during the first deterministic reversal R1. Chemogenetic inhibition of vlOFC increased γd on both reversals (R1 &R3) for female rats, whereas inhibition of BLA increased β on the first probabilistic reversal R3 for female rats. Only comparisons with *$p<0.05$.

834
835 **Figure 6. vlOFC, but not BLA, inhibition impairs adjustment to stimulus-based reversals as**
836 **measured by probability correct.** (**A-B**) Trial structure (A) and timeline (B) of the stimulus-based task.
837 Rats were first surgerized with either hM4Di DREADDs or eGFP null virus on a CaMKII promoter. Rats
838 were allowed to recover for 1 week before testing on a stimulus-or action-based reversal learning task. (**C**)
839 Initial learning of a rewarded stimulus, presented pseudorandomly on the left or right side of the touchscreen
840 for eGFP (top), vlOFC hM4Di (middle), and BLA hM4Di (bottom). (**D**) Plots of accuracy around reversals
841 were restricted to include only animals that reached greater than 50% running window average for the last
842 100 trials in initial discrimination. Rats were always tested on a deterministic schedule before a probabilistic
843 one. Shown are subsequent deterministic (100/0) and probabilistic (90/10) reversal "transitions," 3 sessions
844 before and after each reversal. Drug order was counterbalanced such that on R2 and R4 animals received
845 VEH if they were administered CNO first on R1 and R3, and vice versa. Chemogenetic inhibition of vlOFC
846 abolishes changes in probability correct over the last 3 (pre) and first 3 (post)-reversal sessions, indicating
847 impaired adjustments to reversals. In contrast, BLA inhibition had no clear impact on reversal learning.
848 There was also no effect of CNO in eGFP group learning. **$p<0.01$ different than eGFP following ANOVA
849 of pre-post difference. vlOFC inhibition abolished learning after the first stimulus-based reversal (**Fig. 7**),
850 but did not significantly impair adjustment around action-based probabilistic reversal as following BLA
851 inhibition (**Fig. 8**).
852
853 **Figure 7. BLA inhibition further slows incremental stimulus-based reversal learning whereas vlOFC**
854 **inhibition abolishes learning after first reversal.** Accuracy in stimulus learners measured by mean
855 probability correct for first 5 sessions of each deterministic (100/0) and probabilistic (90/10) reversal. Drug
856 order was counterbalanced such that on R2 and R4 animals received VEH if they were administered CNO
857 first on R1 and R3, and vice versa. Despite animals reaching criterion on initial learning (**Fig. 2C**), animals
858 exhibited poor stimulus-based reversal learning. Reversal learning was particularly flat for R3 (first
859 probabilistic reversal) following vlOFC inhibition compared to eGFP. Reversal learning across all reversals
860 was especially slow for animals following BLA inhibition. Bonferroni-corrected post-hocs following a
861 mixed-effects GLM wherein a session x virus interaction or a session x drug x virus was found resulted in
862 *$p<0.05$ effect of session in eGFP, but not in hM4Di or hM4Di-CNO.

863 **Figure 8. BLA, but not vlOFC, inhibition significantly slowed action-based probabilistic reversal**
864 **adjustment reflected in accuracy around reversals.** Plotted is the probability of choosing the better
865 option (P(correct)) around reversal. Rats were always tested on a deterministic schedule before a
866 probabilistic one. Shown are deterministic (100/0) and probabilistic (90/10) reversal transitions, 100 trials
867 before and after each reversal. Drug order was counterbalanced such that on R2 and R4 animals received

868 VEH if they were administered CNO first on R1 and R3, and vice versa. Inhibition of BLA impaired the
869 first probabilistic reversal learning as indicated by selectively poor performance on R3 reversal. OFC
870 inhibition resulted in no impairment around these reversals. There was also no effect of CNO in the eGFP
871 group.

872

873 **Figure 9. vlOFC, but not BLA, inhibition impairs adjustments to deterministic stimulus-based**
874 **reversals as measured by changes in Win-Stay strategies.** Plots of Win-Stay and Lose-Switch around
875 reversals were restricted to include only animals that reached greater than 50% running window average
876 for the last 100 trials in initial discrimination. **(A)** Inhibition of vlOFC abolishes changes in Win-Stay over
877 the last 3 (pre) and first 3 (post)-reversal sessions, indicating impaired R1-R2 and R2-R3 transitions. In
878 contrast, BLA inhibition had no impact on performance during these transitions. **(B)** Inhibition of BLA
879 attenuates changes in Lose-Switch over the last 3 (pre) and first 3 (post)-reversal sessions in R2-R3, or the
880 transition from deterministic to probabilistic schedule. There was also no effect of CNO in eGFP group
881 learning. **$p<0.01$ different than eGFP following ANOVA of pre-post difference.

882
883
884
885

886 **References**

887 Alsio J, Lehmann O, McKenzie C, Theobald DE, Searle L, Xia J, Dalley JW, Robbins TW
888 (2021) Serotonergic Innervations of the Orbitofrontal and Medial-prefrontal Cortices are
889 Differentially Involved in Visual Discrimination and Reversal Learning in Rats. Cereb
890 Cortex 31:1090-1105.
891 Amodeo LR, McMurray MS, Roitman JD (2017). Orbitofrontal cortex reflects changes in
892 response-outcome contingencies during probabilistic reversal learning. Neuroscience
893 345:27-37.
894 Asrican B, Song J (2021) Extracting meaningful circuit-based calcium dynamics in astrocytes
895 and neurons from adult mouse brain slices using single-photon GCaMP imaging. STAR
896 Protoc 2:100306.
897 Barreiros IV, Panayi MC, Walton ME (2021a) Organization of Afferents along the Anterior-
898 posterior and Medial-lateral Axes of the Rat Orbitofrontal Cortex. Neuroscience 460:53-
899 68.
900 Barreiros IV, Ishii H, Walton ME, Panayi MC (2021b) Defining an orbitofrontal compass:
901 Functional and anatomical heterogeneity across anterior-posterior and medial-lateral
902 axes. Behav Neurosci 135:165-173.
903 Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007) Learning the value of
904 information in an uncertain world. Nat Neurosci 10:1214-1221.
905 Boulougouris V, Robbins TW (2009) Pre-surgical training ameliorates orbitofrontal-mediated
906 impairments in spatial reversal learning. Behavioural Brain Research 197:469-475.
907 Boulougouris V, Dalley JW, Robbins TW (2007) Effects of orbitofrontal, infralimbic and
908 prelimbic cortical lesions on serial spatial reversal learning in the rat. Behavioural Brain
909 Research 179:219-228.
910 Churchwell JC, Morris AM, Heurtelou NM, Kesner RP (2009) Interactions between the
911 prefrontal cortex and amygdala during delay discounting and reversal. Behav Neurosci
912 123:1185-1196.
913 Corbit LH, Balleine BW (2005) Double dissociation of basolateral and central amygdala lesions
914 on the general and outcome-specific forms of pavlovian-instrumental transfer. J Neurosci
915 25:962-970.
916 Costa KM, Scholz R, Lloyd K, Moreno-Castilla P, Gardner MPH, Dayan P, Schoenbaum G
917 (2023) The role of the lateral orbitofrontal cortex in creating cognitive maps. Nat
918 Neurosci 26:107-115.
919 Costa VD, Dal Monte O, Lucas DR, Murray EA, Averbeck BB (2016) Amygdala and Ventral
920 Striatum Make Distinct Contributions to Reinforcement Learning. Neuron 92:505-517.
921 Cox J, Minerva AR, Fleming WT, Zimmerman CA, Hayes C, Zorowitz S, Bandi A, Ornelas S,
922 McMannon B, Parker NF, Witten IB (2023) A neural substrate of sex-dependent
923 modulation of motivation. Nat Neurosci 26(2):274-284.
924 Dalton GL, Wang NY, Phillips AG, Floresco SB (2016) Multifaceted Contributions by Different
925 Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal
926 Learning. J Neurosci 36:1996-2006.
927 Esber GR, Roesch MR, Bali S, Trageser J, Bissonette GB, Puche AC, Holland PC, Schoenbaum
928 G (2012) Attention-related Pearce-Kaye-Hall signals in basolateral amygdala require the
929 midbrain dopaminergic system. Biol Psychiatry 72:1012-1019.
930 Ghasemi A, Jeddi S, Kashfi K (2021) The laboratory rat: Age and body weight matter. EXCLI J
931 20:1431-1445.

932 Grissom NM, Reyes TM (2019) Let's call the whole thing off: evaluating gender and sex
933 differences in executive function. Neuropsychopharmacology 44:86-96.
934 Groman SM, Keistler C, Keip AJ, Hammarlund E, DiLeone RJ, Pittenger C, Lee D, Taylor JR
935 (2019) Orbitofrontal Circuits Control Multiple Reinforcement-Learning Processes.
936 Neuron 103:734-746 e733.
937 Harris C, Aguirre C, Kolli S, Das K, Izquierdo A, Soltani A (2021) Unique features of stimulus-
938 based probabilistic reversal learning. Behav Neurosci 135:550-570.
939 Hart EE, Blair GJ, O'Dell TJ, Blair HT, Izquierdo A (2020) Chemogenetic Modulation and
940 Single-Photon Calcium Imaging in Anterior Cingulate Cortex Reveal a Mechanism for
941 Effort-Based Decisions. J Neurosci 40:5628-5643.
942 Hervig ME, Fiddian L, Piilgaard L, Bozic T, Blanco-Pozo M, Knudsen C, Olesen SF, Alsio J,
943 Robbins TW (2020) Dissociable and Paradoxical Roles of Rat Medial and Lateral
944 Orbitofrontal Cortex in Visual Serial Reversal Learning. Cereb Cortex 30:1016-1029.
945 Iordanova MD, Yau JO-Y, McDannald MA, Corbit LH (2021) Neural substrates of appetitive
946 and aversive prediction error. Neurosci Biobehav Rev 123:337-351.
947 Izquierdo A (2017) Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward
948 Learning and Decision Making. J Neurosci 37:10529-10540.
949 Izquierdo A, Murray EA (2004) Combined Unilateral Lesions of the Amygdala and Orbital
950 Prefrontal Cortex Impair Affective Processing in Rhesus Monkeys. Journal of
951 Neurophysiology 91:2023-2039.
952 Izquierdo A, Darling C, Manos N, Pozos H, Kim C, Ostrander S, Cazares V, Stepp H, Rudebeck
953 PH (2013) Basolateral amygdala lesions facilitate reward choices after negative feedback
954 in rats. J Neurosci 33:4105-4109.
955 Janak PH, Tye KM (2015) From circuits to behaviour in the amygdala. Nature 517:284-292.
956 Jang AI, Costa VD, Rudebeck PH, Chudasama Y, Murray EA, Averbeck BB (2015) The Role of
957 Frontal Cortical and Medial-Temporal Lobe Brain Areas in Learning a Bayesian Prior
958 Belief on Reversals. J Neurosci 35:11751-11760.
959 Jenni NL, Rutledge G, Floresco SB (2022) Distinct Medial Orbitofrontal-Striatal Circuits
960 Support Dissociable Component Processes of Risk/Reward Decision-Making. J Neurosci
961 42:2743-2755.
962 Jones B, Mishkin M (1972). Limbic lesions and the problem of stimulus-reinforcement
963 associations. Experimental Neurology 36:362–377.
964 Keefer SE, Petrovich GD (2022) Necessity and recruitment of cue-specific neuronal ensembles
965 within the basolateral amygdala during appetitive reversal learning. Neurobiol Learn
966 Mem 194:107663.
967 Kim J, Ragozzino ME (2005) The involvement of the orbitofrontal cortex in learning under
968 changing task contingencies. Neurobiol Learn Mem 83:125-133.
969 Lichtenberg NT, Pennington ZT, Holley SM, Greenfield VY, Cepeda C, Levine MS, Wassum
970 KM (2017) Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-
971 Triggered Reward Expectations. J Neurosci 37:8374-8384.
972 Malvaez M, Shieh C, Murphy MD, Greenfield VY, Wassum KM (2019) Distinct cortical-
973 amygdala projections drive reward value encoding and retrieval. Nat Neurosci 22:762-
974 769.
975 Namboodiri VMK, Hobbs T, Trujillo-Pisanty I, Simon RC, Gray MM, Stuber GD (2021)
976 Relative salience signaling within a thalamo-orbitofrontal circuit governs learning rate.
977 Curr Biol 31:5176-5191.e5175.

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.

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 Processes 200:104663-104663.

Pachitariu M, Stringer C, Dipoppa M, Schröder S, Rossi LF, Dalgleish 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, 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.

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.

Sias AC, Morse AK, Wang S, Greenfield VY, Goodpaster CM, Wrenn TM, Wikenheiser AM, Holley SM, Cepeda C, Levine MS, Wassum KM (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.

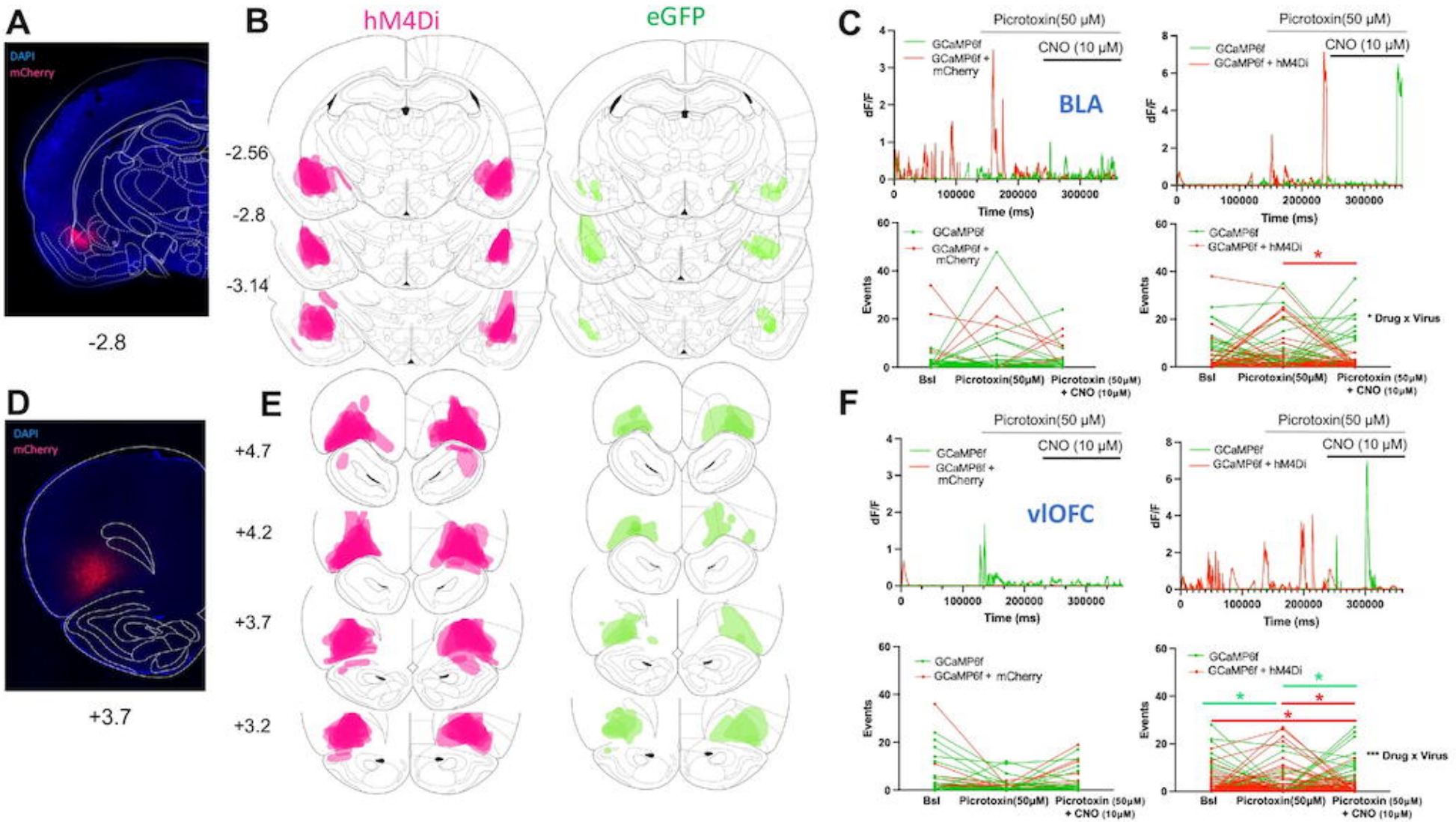
Soltani A, Koechlin E (2022) Computational models of adaptive behavior and prefrontal cortex. Neuropsychopharmacology 47: 58-71.

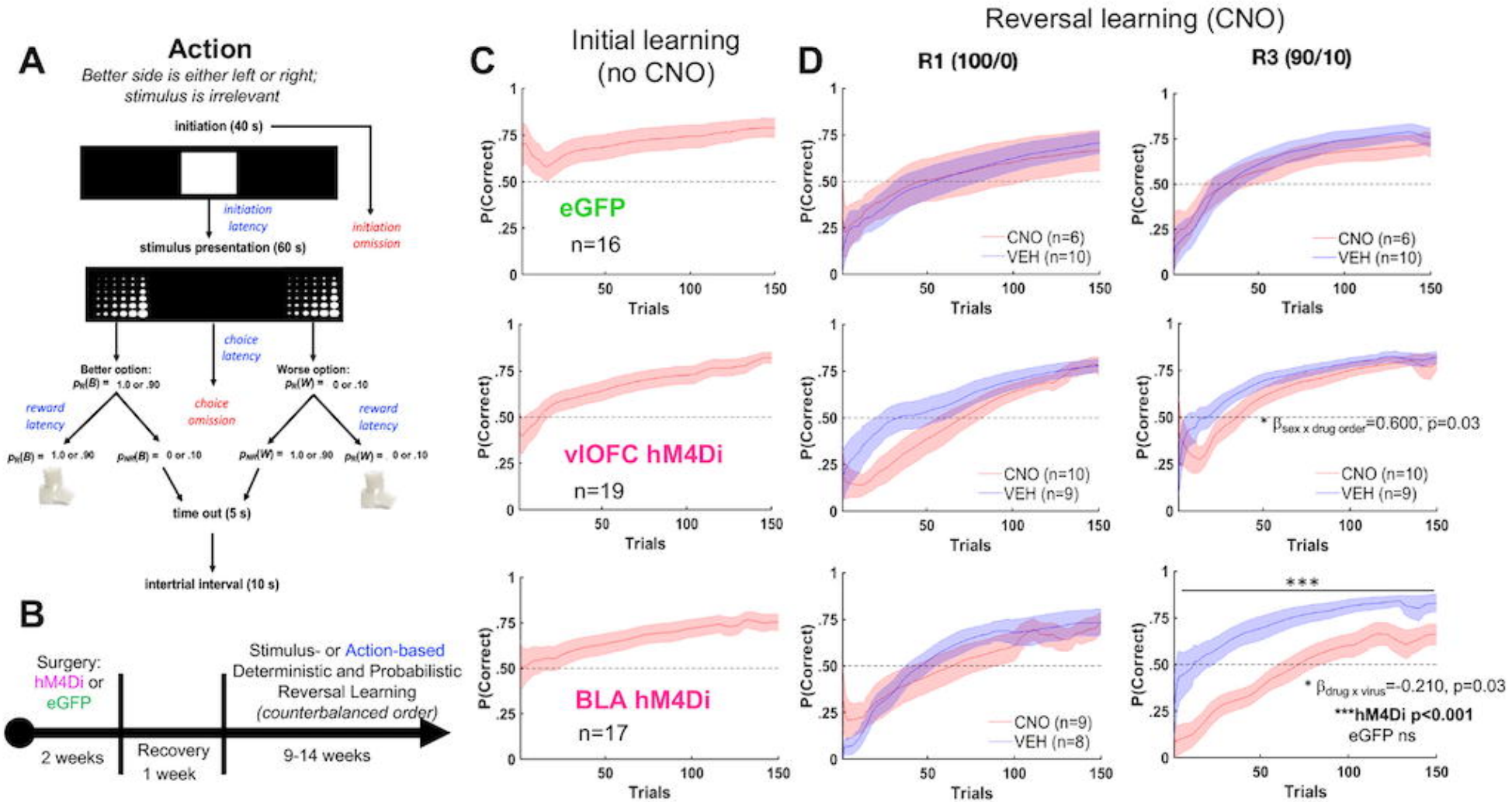
Stalnaker TA, Roesch MR, Calu DJ, Burke KA, Singh T, Schoenbaum G (2007) Neural correlates of inflexible behavior in the orbitofrontal-amygdalar circuit after cocaine exposure. Ann N Y Acad Sci 1121:598-609.

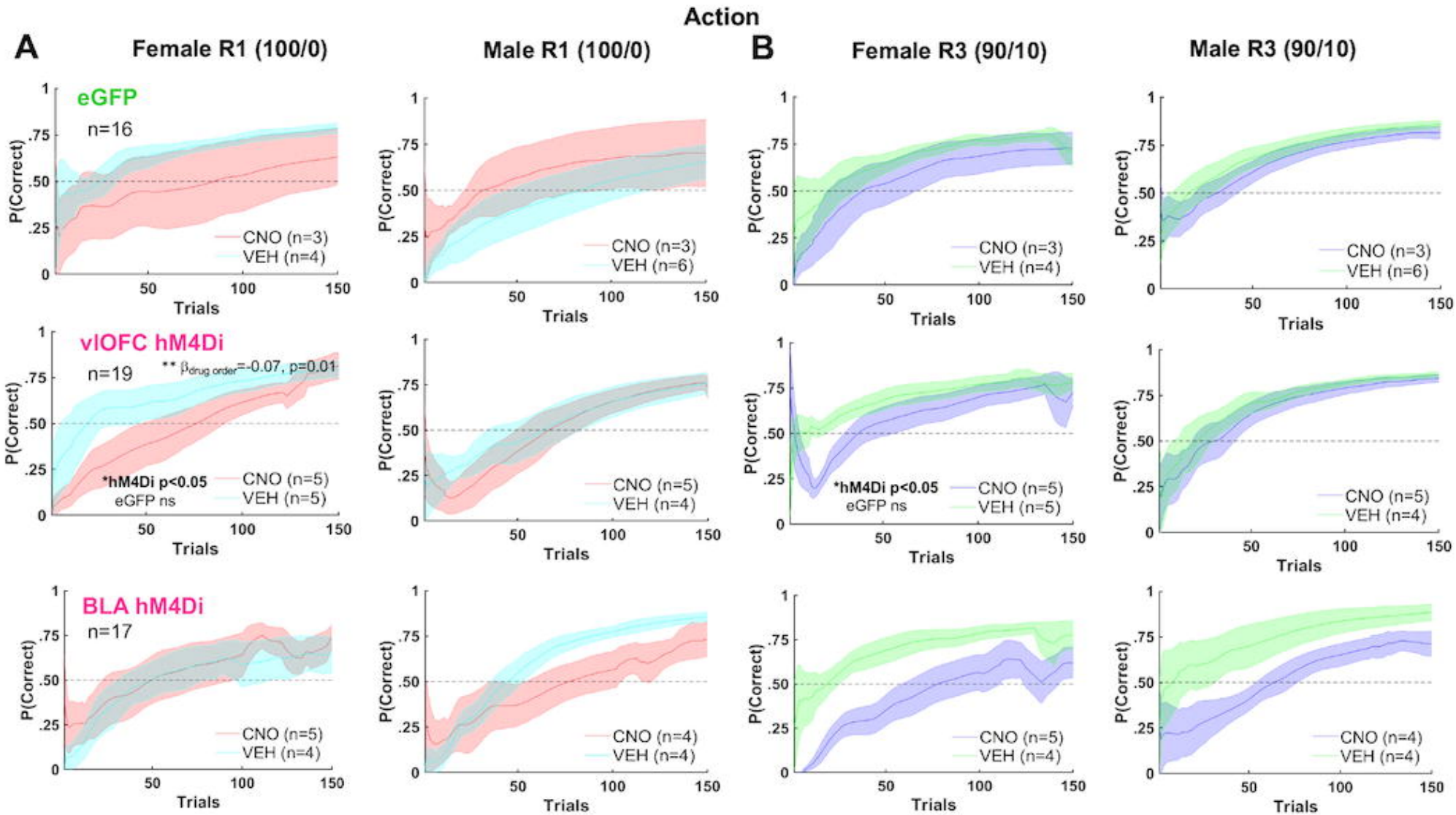
Stolyarova A, Izquierdo A (2017) Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. Elife 6.

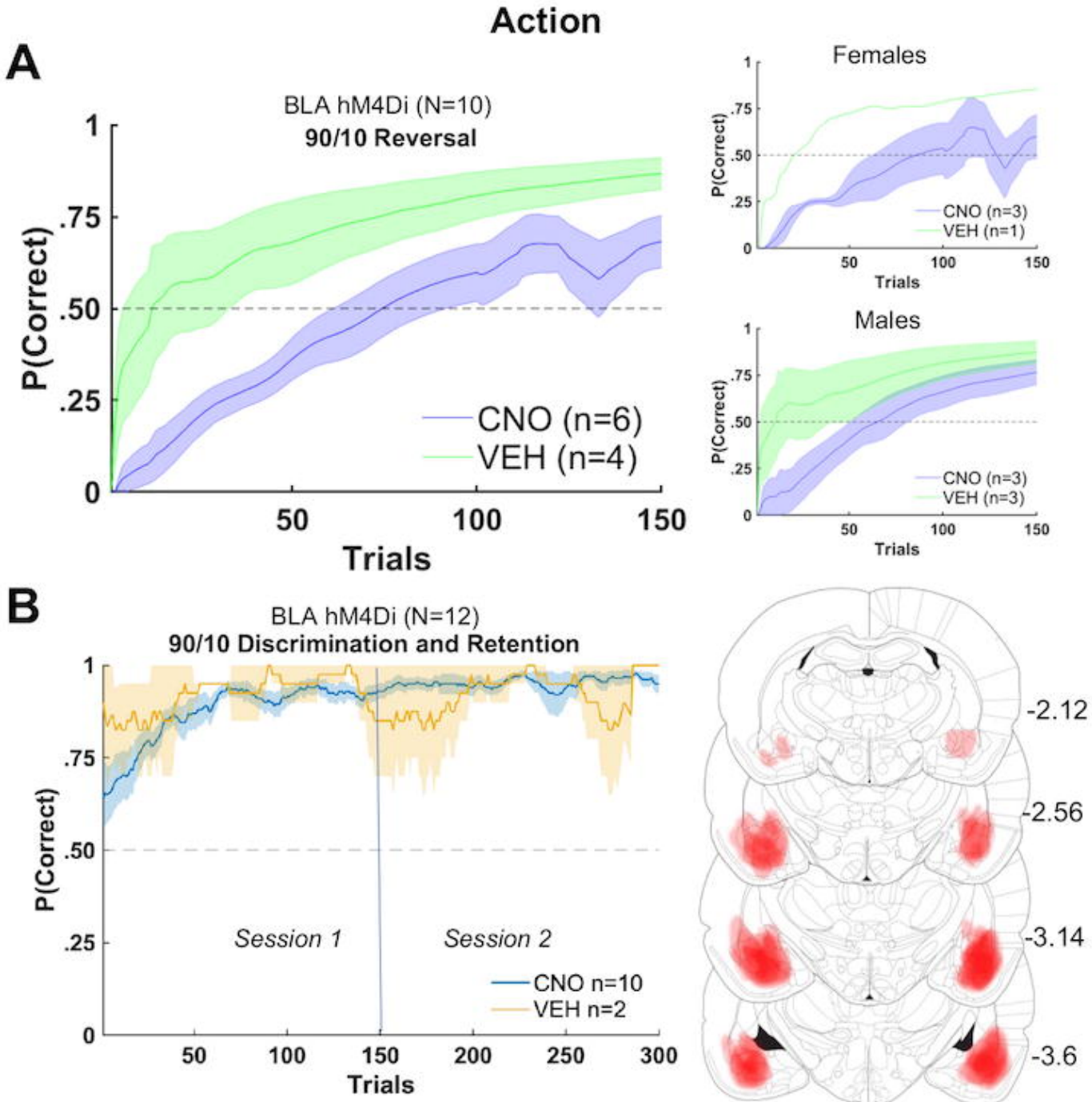
Stolyarova A, Rakhshan M, Hart EE, O'Dell TJ, Peters MAK, Lau H, Soltani A, Izquierdo A (2019) Contributions of anterior cingulate cortex and basolateral amygdala to decision confidence and learning under uncertainty. Nat Commun 10:4704.

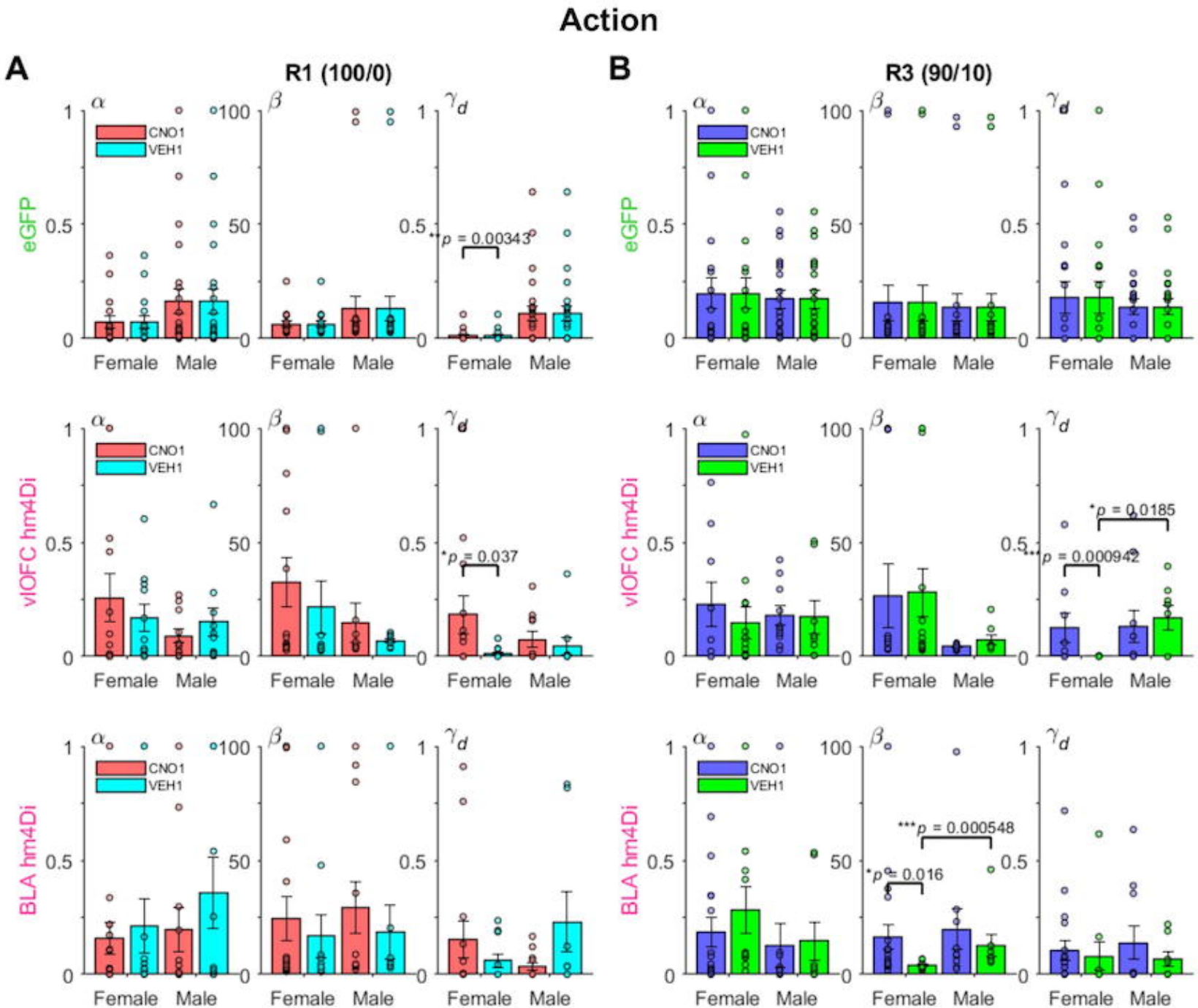
1024 Tada M, Takeuchi A, Hashizume M, Kitamura K, Kano M (2014) A highly sensitive fluorescent
1025 indicator dye for calcium imaging of neural activity in vitro and in vivo. Eur J Neurosci
1026 39:1720-1728.
1027 Taswell CA, Costa VD, Basile BM, Pujara MS, Jones B, Manem N, Murray EA, Averbeck BB
1028 (2021) Effects of Amygdala Lesions on Object-Based Versus Action-Based Learning in
1029 Macaques. Cereb Cortex 31:529-546.
1030 Tye KM, Janak PH (2007) Amygdala neurons differentially encode motivation and
1031 reinforcement. J Neurosci 27:3937-3945.
1032 Verharen JPH, den Ouden HEM, Adan RAH, Vanderschuren L (2020) Modulation of value-
1033 based decision making behavior by subregions of the rat prefrontal cortex.
1034 Psychopharmacology (Berl) 237:1267-1280.
1035 Ward RD, Winiger V, Kandel ER, Balsam PD, Simpson EH (2015) Orbitofrontal cortex
1036 mediates the differential impact of signaled-reward probability on discrimination
1037 accuracy. Front Neurosci 9:230.
1038 Wassum KM, Izquierdo A (2015) The basolateral amygdala in reward learning and addiction.
1039 Neurosci Biobehav Rev 57:271-283.
1040 Winstanley CA, Floresco SB (2016) Deciphering Decision Making: Variation in Animal Models
1041 of Effort- and Uncertainty-Based Choice Reveals Distinct Neural Circuitries Underlying
1042 Core Cognitive Processes. J Neurosci 36:12069-12079.
1043 Wright SL, Martin GM, Thorpe CM, Haley K, Skinner DM (2019) Distance and direction, but
1044 not light cues, support response reversal learning. Learning & Behavior 47:38-46.
1045 Ye T, Romero-Sosa JL, Rickard A, Aguirre CG, Wikenheiser AM, Blair HT, Izquierdo A (2023)
1046 Theta oscillations in anterior cingulate cortex and orbitofrontal cortex differentially
1047 modulate accuracy and speed in flexible reward learning. Oxford Open Neuroscience
1048 kvad005.
1049 Zimmermann KS, Li CC, Rainnie DG, Ressler KJ, Gourley SL (2018) Memory Retention
1050 Involves the Ventrolateral Orbitofrontal Cortex: Comparison with the Basolateral
1051 Amygdala. Neuropsychopharmacology 43:373-383.
1052

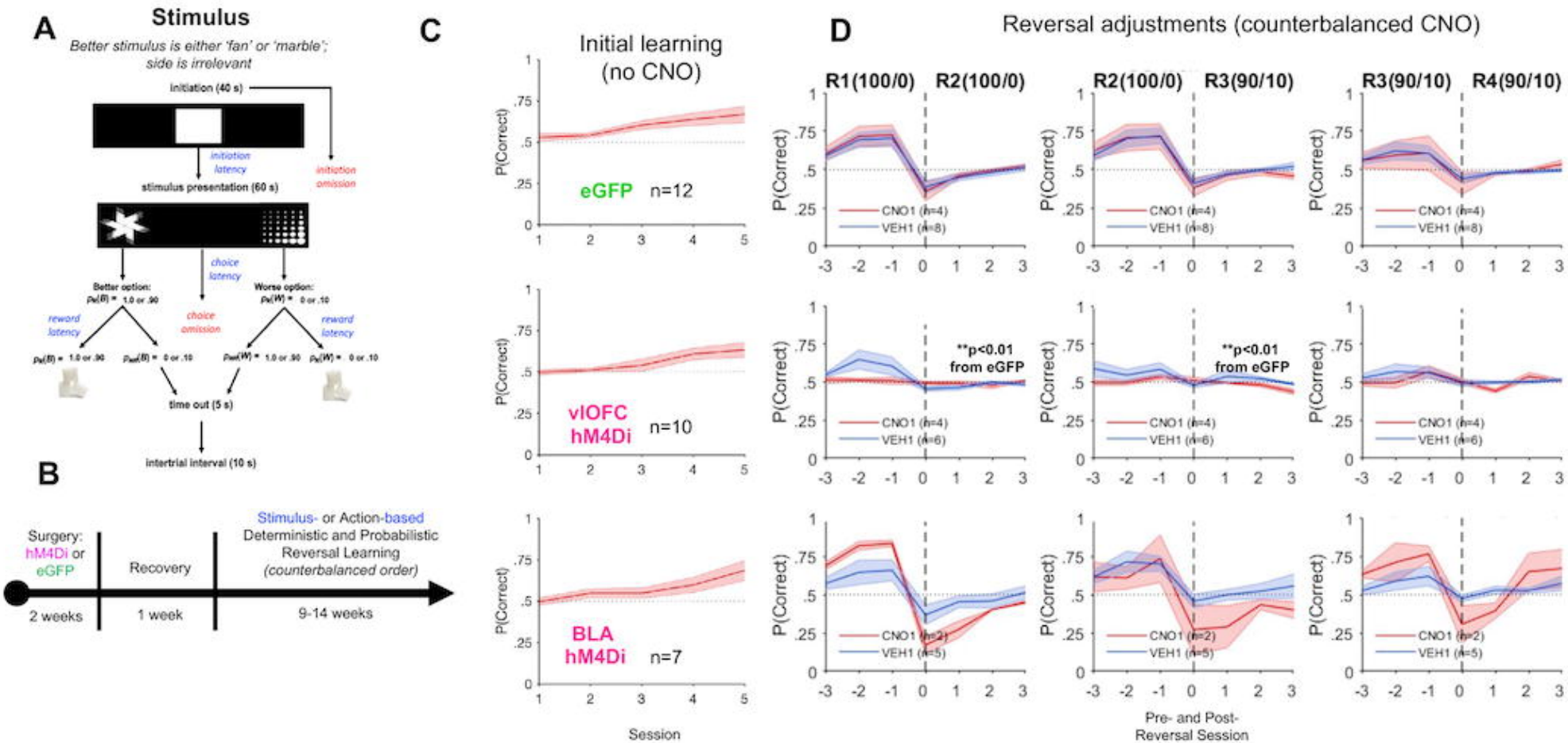

A
DAPI
mCherry
-2.8
B
hM4Di
eGFP
-2.56
-2.8
-3.14
C
Picrotoxin(50 μM)
CNO (10 μM)
GCaMP6f
GCaMP6f + mCherry
GCaMP6f + hM4Di
BLA
dF/F
Time (ms)
Events
Bsl
Picrotoxin(50μM)
Picrotoxin (50μM) + CNO (10μM)
* Drug x Virus
D
DAPI
mCherry
+3.7
E
+4.7
+4.2
+3.7
+3.2
F
vlOFC
*** Drug x Virus

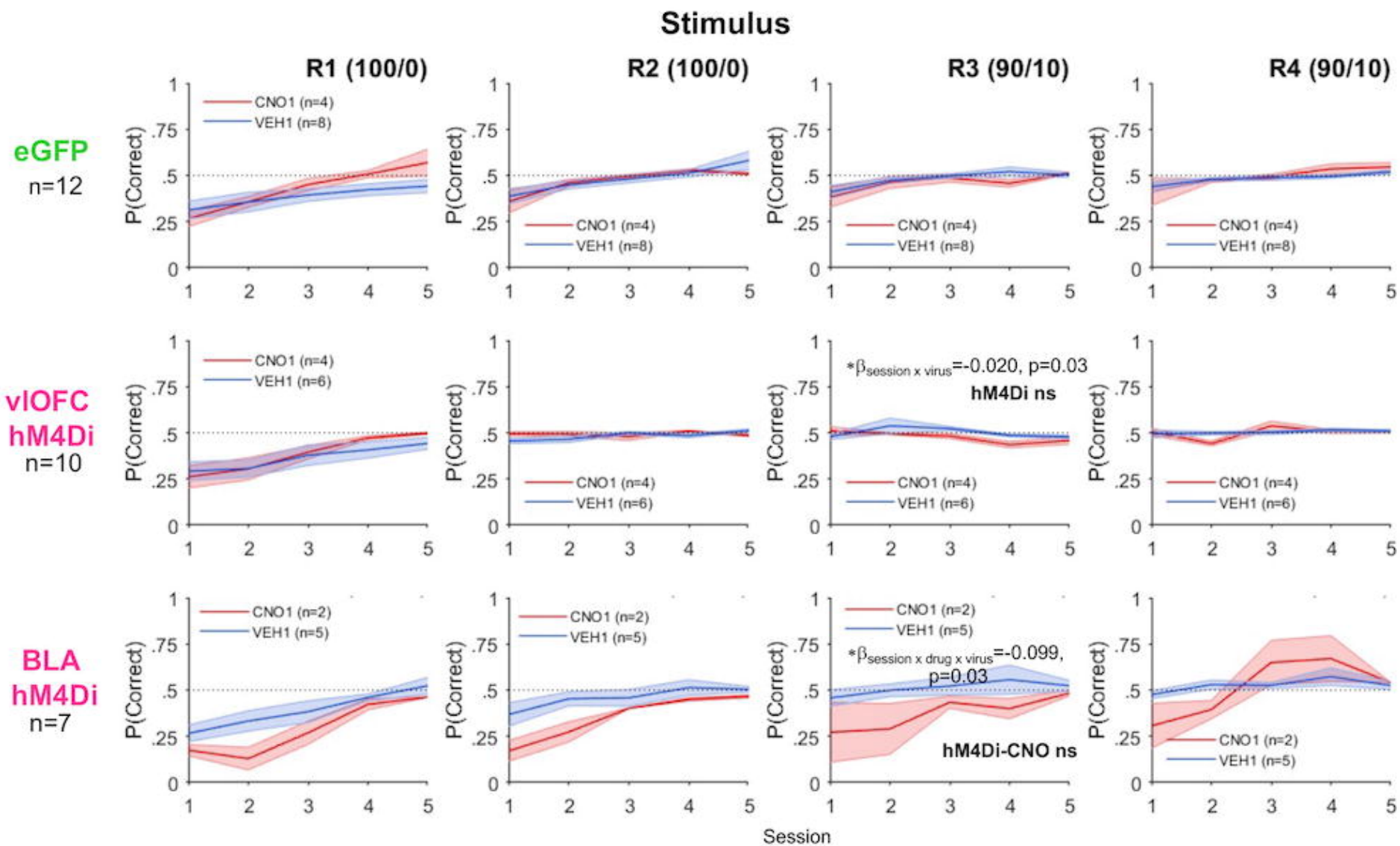

A
Action
Better side is either left or right;
stimulus is irrelevant
initiation (40 s)
initiation latency
initiation omission
stimulus presentation (60 s)
choice latency
Better option:
Worse option:
reward latency
choice omission
reward latency
time out (5 s)
intertrial interval (10 s)
B
Surgery:
hM4Di or
eGFP
2 weeks
Recovery
1 week
Stimulus- or Action-based
Deterministic and Probabilistic
Reversal Learning
(counterbalanced order)
9-14 weeks
C
Initial learning
(no CNO)
eGFP
n=16
vlOFC hM4Di
n=19
BLA hM4Di
n=17
P(Correct)
Trials
D
Reversal learning (CNO)
R1 (100/0)
R3 (90/10)
CNO (n=6)
VEH (n=10)
CNO (n=10)
VEH (n=9)
CNO (n=9)
VEH (n=8)
* β sex x drug order=0.600, p=0.03
* β drug x virus=-0.210, p=0.03
***hM4Di p<0.001
eGFP ns

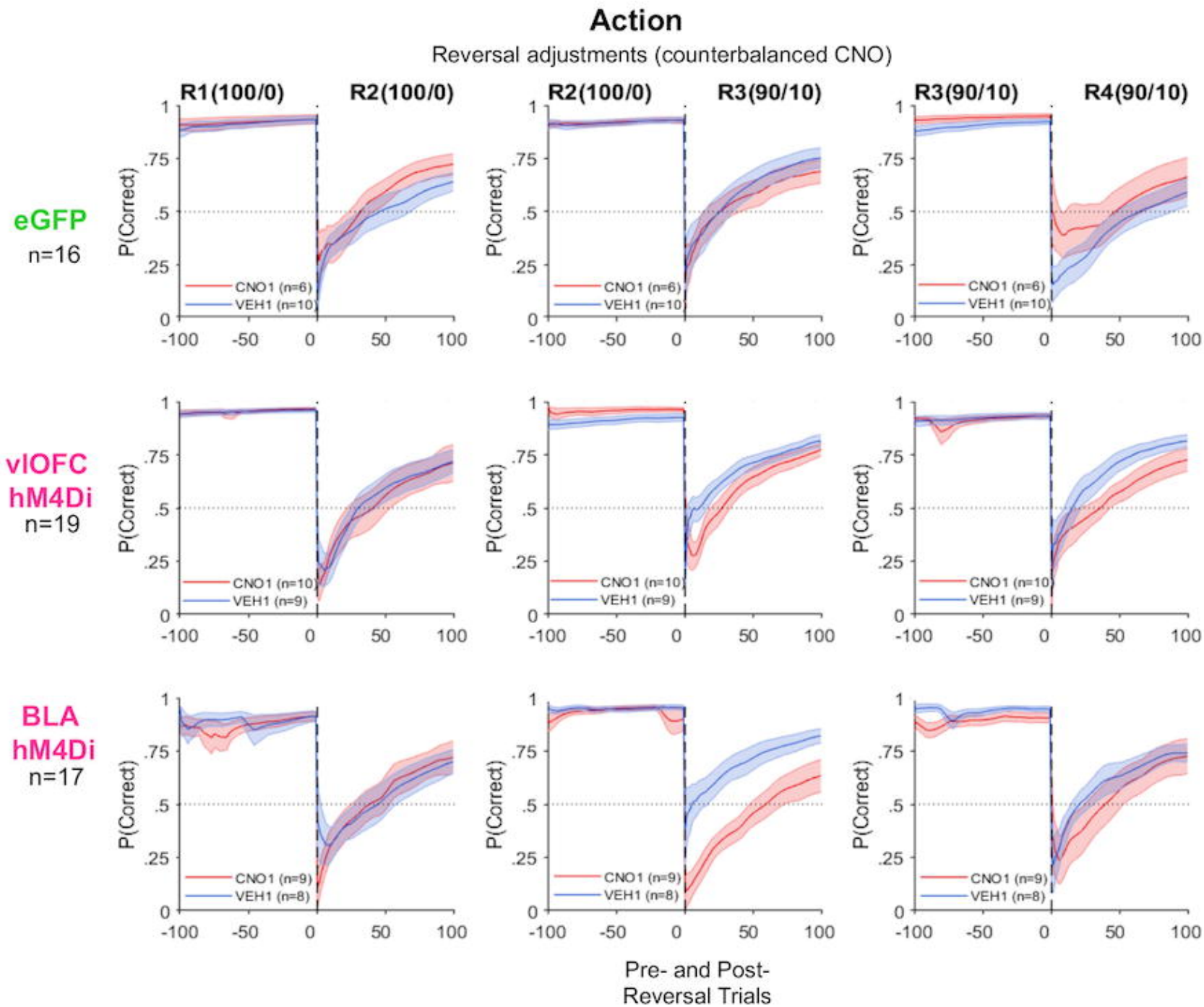

Action
A
Female R1 (100/0)
Male R1 (100/0)
B
Female R3 (90/10)
Male R3 (90/10)
eGFP
n=16
CNO (n=3)
VEH (n=4)
CNO (n=3)
VEH (n=6)
CNO (n=3)
VEH (n=4)
CNO (n=3)
VEH (n=6)
vlOFC hM4Di
n=19
** $\beta_{drug\ order}$=-0.07, p=0.01
*hM4Di p<0.05
eGFP ns
CNO (n=5)
VEH (n=5)
CNO (n=5)
VEH (n=4)
*hM4Di p<0.05
eGFP ns
CNO (n=5)
VEH (n=5)
CNO (n=5)
VEH (n=4)
BLA hM4Di
n=17
CNO (n=5)
VEH (n=4)
CNO (n=4)
VEH (n=4)
CNO (n=5)
VEH (n=4)
CNO (n=4)
VEH (n=4)
P(Correct)
Trials
0
.25
.50
.75
1
50
100
150

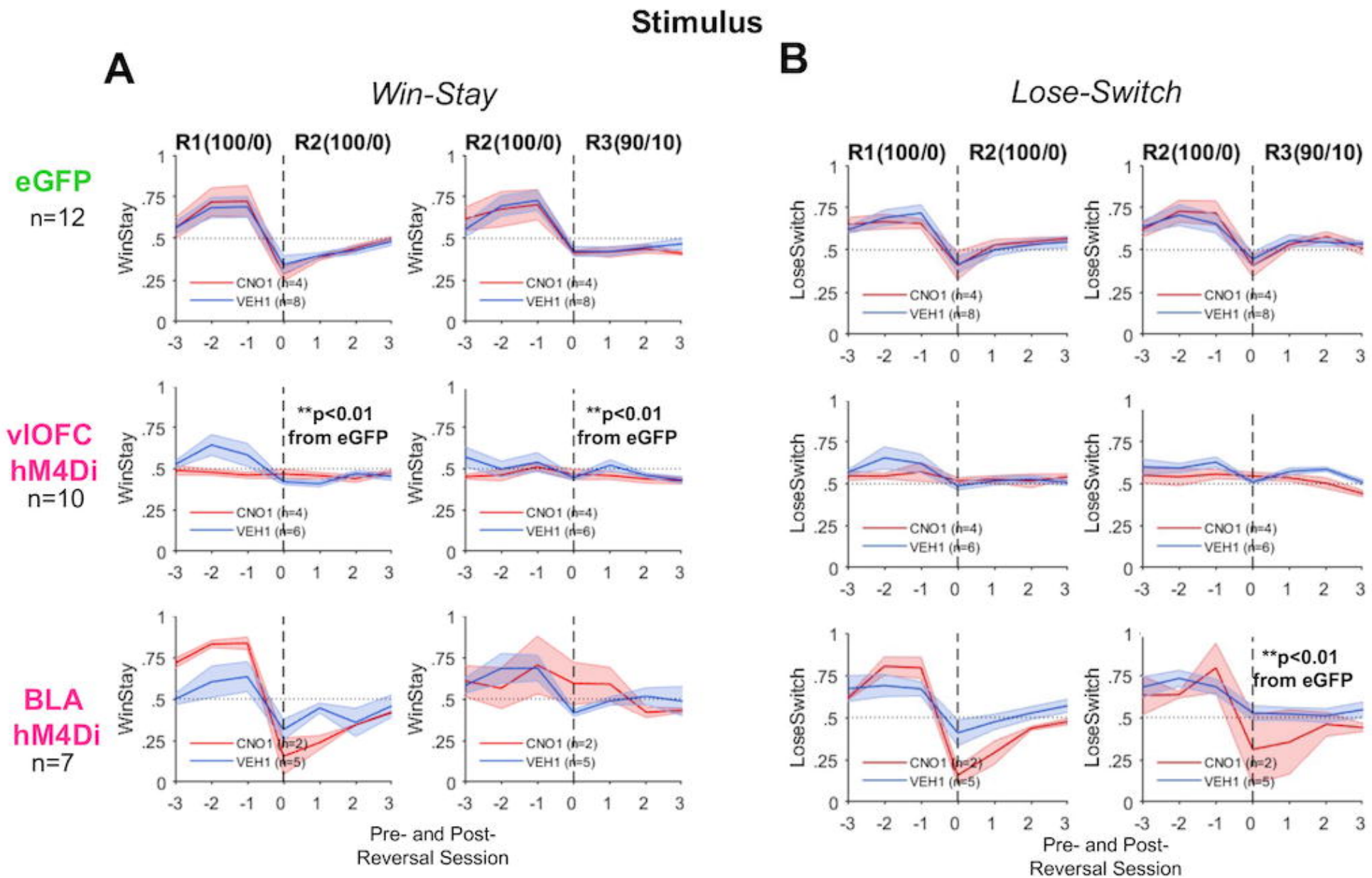

Action
A
BLA hM4Di (N=10)
90/10 Reversal
P(Correct)
Trials
CNO (n=6)
VEH (n=4)
Females
CNO (n=3)
VEH (n=1)
Males
CNO (n=3)
VEH (n=3)
B
BLA hM4Di (N=12)
90/10 Discrimination and Retention
Session 1
Session 2
CNO n=10
VEH n=2
-2.12
-2.56
-3.14
-3.6


Action
A
R1 (100/0)
B
R3 (90/10)
eGFP
vlOFC hm4Di
BLA hm4Di
CNO1
VEH1
Female
Male
**p = 0.00343
*p = 0.037
*p = 0.0185
**p = 0.000942
***p = 0.000548
*p = 0.016


A
Stimulus
Better stimulus is either 'fan' or 'marble'; side is irrelevant
initiation (40 s)
initiation latency
initiation omission
stimulus presentation (60 s)
choice latency
Better option:
Worse option:
reward latency
choice omission
reward latency
time out (5 s)
intertrial interval (10 s)
B
Surgery: hM4Di or eGFP
Recovery
Stimulus- or Action-based Deterministic and Probabilistic Reversal Learning (counterbalanced order)
2 weeks
1 week
9-14 weeks
C
Initial learning (no CNO)
eGFP n=12
vlOFC hM4Di n=10
BLA hM4Di n=7
P(Correct)
Session
D
Reversal adjustments (counterbalanced CNO)
R1(100/0) R2(100/0)
R2(100/0) R3(90/10)
R3(90/10) R4(90/10)
CNO1 (n=4)
VEH1 (n=8)
CNO1 (n=4)
VEH1 (n=6)
CNO1 (n=2)
VEH1 (n=5)
**p<0.01 from eGFP
Pre- and Post-Reversal Session


Stimulus
R1 (100/0)
R2 (100/0)
R3 (90/10)
R4 (90/10)
eGFP
n=12
vlOFC
hM4Di
n=10
BLA
hM4Di
n=7
P(Correct)
CNO1 (n=4)
VEH1 (n=8)
CNO1 (n=4)
VEH1 (n=6)
CNO1 (n=2)
VEH1 (n=5)
*βsession x virus=-0.020, p=0.03
hM4Di ns
*βsession x drug x virus=-0.099, p=0.03
hM4Di-CNO ns
Session


Action
Reversal adjustments (counterbalanced CNO)
R1(100/0)
R2(100/0)
R2(100/0)
R3(90/10)
R3(90/10)
R4(90/10)
eGFP
n=16
vlOFC
hM4Di
n=19
BLA
hM4Di
n=17
P(Correct)
CNO1 (n=6)
VEH1 (n=10)
CNO1 (n=10)
VEH1 (n=9)
CNO1 (n=9)
VEH1 (n=8)
Pre- and Post-
Reversal Trials


Stimulus
A
Win-Stay
B
Lose-Switch
R1(100/0)
R2(100/0)
R3(90/10)
eGFP
n=12
vlOFC
hM4Di
n=10
BLA
hM4Di
n=7
WinStay
LoseSwitch
CNO1 (n=4)
VEH1 (n=8)
CNO1 (n=4)
VEH1 (n=6)
CNO1 (n=2)
VEH1 (n=5)
**p<0.01
from eGFP
Pre- and Post-
Reversal Session