Theta oscillations in anterior cingulate cortex and orbitofrontal cortex differentially modulate accuracy and speed in flexible reward learning

Tony Ye[1], Juan Luis Romero-Sosa[1], Anne Rickard[1], Claudia G. Aguirre[1], Andrew M. Wikenheiser[1,2,3,4], Hugh T. Blair[1,2,3] and Alicia Izquierdo[1,2,3,4,*,†]

[1]Department of Psychology, UCLA, Los Angeles, CA 90095, USA
[2]The Brain Research Institute, UCLA, Los Angeles, CA 90095, USA
[3]Integrative Center for Learning and Memory, UCLA, Los Angeles, CA 90095, USA
[4]Integrative Center for Addictions, UCLA, Los Angeles, CA 90095, USA
*Correspondence address. Tel: +1 310 825 3459; E-mail: aizquie@psych.ucla.edu
†Alicia Izquierdo is a senior editor of *Oxford Open Neuroscience*. Full peer-review for this manuscript was handled by *Oxford Open Neuroscience* Senior Editor Sam Gilbert.

Abstract

Flexible reward learning relies on frontal cortex, with substantial evidence indicating that anterior cingulate cortex (ACC) and orbitofrontal cortex (OFC) subregions play important roles. Recent studies in both rat and macaque suggest theta oscillations (5–10 Hz) may be a spectral signature that coordinates this learning. However, network-level interactions between ACC and OFC in flexible learning remain unclear. We investigated the learning of stimulus–reward associations using a combination of simultaneous *in vivo* electrophysiology in dorsal ACC and ventral OFC, partnered with bilateral inhibitory DREADDs in ACC. In freely behaving male and female rats and using a within-subject design, we examined accuracy and speed of response across distinct and precisely defined trial epochs during initial visual discrimination learning and subsequent reversal of stimulus–reward contingencies. Following ACC inhibition, there was a propensity for random responding in early reversal learning, with correct vs. incorrect trials distinguished only from OFC, not ACC, theta power differences in the reversal phase. ACC inhibition also hastened incorrect choices during reversal. This same pattern of change in accuracy and speed was not observed in viral control animals. Thus, characteristics of impaired reversal learning following ACC inhibition are poor deliberation and weak theta signaling of accuracy in this region. The present results also point to OFC theta oscillations as a prominent feature of reversal learning, unperturbed by ACC inhibition.

Keywords: DREADDs., theta, prefrontal cortex, oscillations, reversal learning

INTRODUCTION

Both humans and nonhumans alike make a multitude of choices based on adaptations to their environment in order to maximize rewards while minimizing losses. This type of flexible reward learning is thought to rely on the prefrontal cortex. Substantial evidence in primates indicates that anterior cingulate cortex (ACC) and orbitofrontal cortex (OFC) have essential roles in learning about actions and stimuli, respectively [1–3]. Although primate and rodent ACC and OFC are not equivalent structures, we may still learn fundamental principles of their functions in the rodent [4]. A mechanistic investigation of their interaction in stimulus-based learning is missing in the broader literature.

Lesion or inhibition of OFC impairs multiple forms of reward learning and in rats this has most extensively been observed in the reversal phase [5, 6], particularly in spatial- or action-based reversal learning [7–10]. Classic recording studies in rat OFC during (olfactory cue) reversal learning [11] show that neural activity represents an integration of the motivational significance of a reward and its specific predictive cue, critical

for reversal but not discrimination learning [11, 12]. Conversely, lesion or inhibition of ACC results in decrements in (visual) discrimination performance [13], but has no effect on fully predictive, deterministic reversal learning [14, 15]. Yet ACC neurons in rats integrate motor-related information with expectations of future outcomes [16, 17], encode choice value [18] and maintain trial-by-trial performance [19], all processes that would be vital not just to action-based learning but also to stimulus-based reversal learning. This growing body of evidence supports a general role for ACC in maintaining a model about the reward environment and making predictions about future choices associated with either actions or stimuli.

More importantly, how these two regions interact is unclear. Several groups have recorded single-unit and local-field activity simultaneously in both areas and uncovered overlapping functions, in contrast to the interference studies [3, 20]. For example, there is reported evidence of ACC-OFC task-related coherence in theta to low beta (4–20 Hz) oscillations as rats respond to or choose a higher-valued reward, suggesting a functional
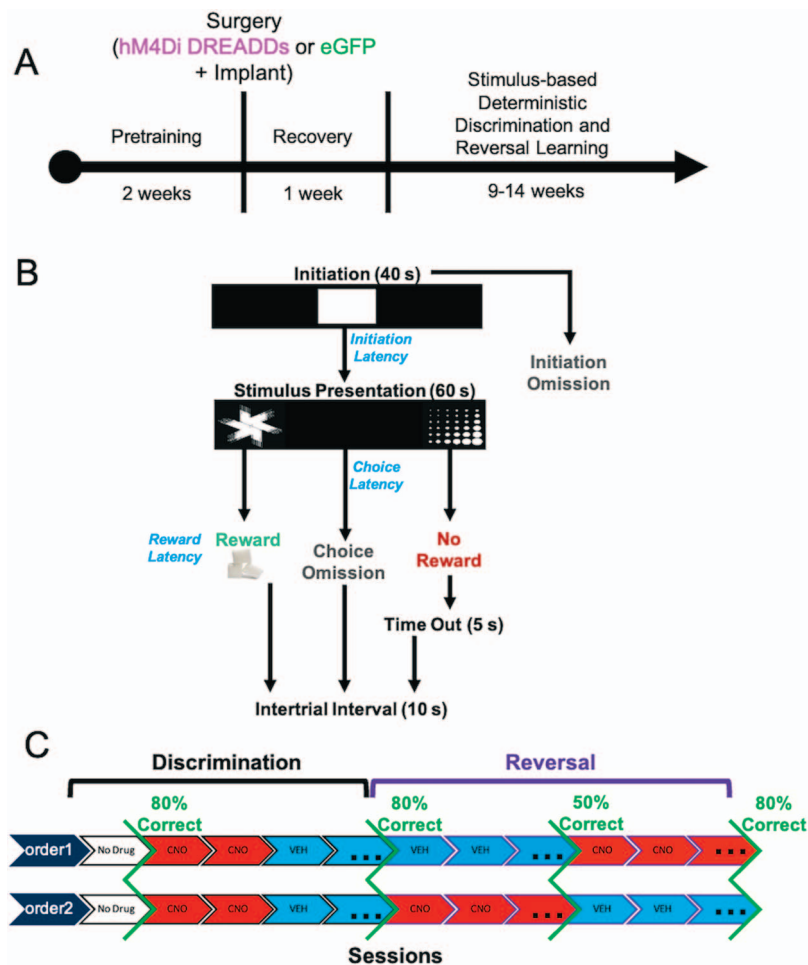
Figure 1. Experimental timeline, task structure, and drug administration. (**A**) After approximately 2 weeks of pretraining, rats underwent bilateral viral infusion of inhibitory hM4Di DREADDs or null virus enhanced green fluorescent protein (eGFP) in ACC along with chronic implantation of electrode arrays targeting both ACC and OFC unilaterally in each animal. (**B**) Following a 1-week recovery period, rats commenced the stimulus-based visual discrimination and reversal learning task, with rewarded-stimulus assignment (e.g. 'fan' vs. 'marbles') counterbalanced across rats. (**C**) Drug administration (clozapine-N-oxide, CNO or vehicle, VEH) was administered once criterion was reached for initial discrimination to probe the effect of ACC inhibition on mastery-level discrimination performance. Subsequently, in the reversal phase, the order of drug administration was counterbalanced across animals.

connectivity underlying relative reward value [21–23]. It is therefore conceivable that these two regions work in a complementary manner for reward-based learning and choice [24, 25] potentially with theta oscillations in OFC coordinating reward signals [12, 26] with ACC [23]. However, network-level interactions and the spectral signatures in flexible reward learning remain unclear.

Given that ACC is densely innervated by regions involved not only in evaluating options but also preparing and executing actions, we hypothesized that ACC theta modulates overall learning of stimuli. To test this, we investigated learning using a combination of simultaneous *in-vivo* electrophysiology in dorsal ACC (dorsal area 32/24, van Heukelum, Mars et al. [27]) and ventral OFC, partnered with inhibitory Designer Receptors Exclusively Activated by Designer Drugs (DREADDs) in ACC. We recorded from posterior VO as it is richly interconnected with ACC [28, 29]. Additionally, a thorough consideration of both accuracy and speed of response during learning may be particularly informative when identifying neural signatures of flexible learning [30, 31], so we include analysis of both here.

We found that ACC inhibition during the discrimination phase had no functional impact on initial performance. In contrast, ACC inhibition in early reversal learning promoted a random response strategy, disrupted ACC signaling of correct vs. incorrect trials, and hastened speed of incorrect choices. The same pattern was not observed in viral control animals. Although ACC inhibition produced these changes in accuracy and speed in the reversal phase, it did not affect OFC theta signaling of accuracy in reversal learning, suggesting an independent process in OFC.

RESULTS
ACC inhibition in early, not late, reversal learning promotes random responding

Fifteen rats were prepared with either inhibitory (hM4Di) DREADDs (n = 8, 4 females) or enhanced Green Fluorescent Protein (eGFP) null virus (n = 7, 3 females) on putative projection neurons in ACC. Eight of these animals (n = 4 hM4Di, n = 4 eGFP) were also implanted with custom-constructed 16-channel fixed electrode arrays unilaterally in ACC and OFC, in the same surgery as virus infusion. All rats were tested using a within-subject design for drug administration (Fig. 1). Rats were administered clozapine-N-oxide (CNO) once they reached criterion-level performance (80%) on the initial discrimination learning phase to assess the impact of ACC inhibition on mastery-level performance [13, 14].
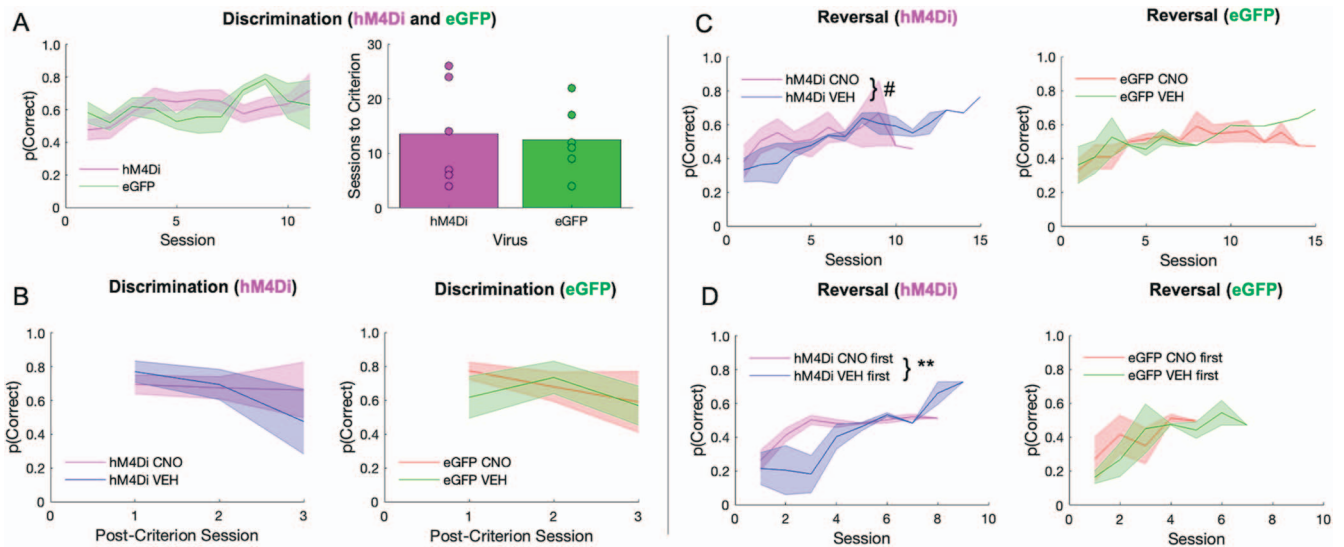
Figure 2. ACC inhibition has no effect on discrimination performance, but promotes random responding during early reversal learning. There were no differences between hM4Di and eGFP virus groups in the initial discrimination phase, however, clozapine-N-oxide (CNO) affected long-term reversal learning differently depending on virus group, and depending on when CNO was administered. **(A)** Performance in hM4Di rats and eGFP rats during initial discrimination of stimuli during which there was no recording and no drug administered (left). Rats did not differ by virus group in the number of sessions to reach criterion (right). **(B)** There was no difference between virus groups or by drug in post-criterion discrimination performance. **(C)** Learning curves during the reversal phase split by hM4Di (left) and eGFP (right). Irrespective of virus condition, rats rarely surpassed 60% accuracy. **(D)** Rats received either VEH or CNO first during the reversal phase prior to reaching 50% correct. Same learning as in C, but learning curves now split by whether rats received CNO first or VEH first. ACC inhibition during early reversal learning promotes a random response strategy that they do not recover from. Means ± standard error of the mean (S.E.M.) shades. Post-hoc Bonferroni comparisons, #$P = 0.07$, **$P = 0.001$. hM4Di $n = 8$ (4 female), and eGFP $n = 7$ (3 female).

Subsequently in the reversal phase, they were administered CNO or vehicle (VEH) in counterbalanced order: CNO administered early in learning (to 50% correct) then switched to VEH until 80%, or vice versa (VEH to 50%, then CNO to 80%). We used a CNO dose for ACC inhibition that has previously resulted in behavioral effects on two different reward learning tasks [32, 33], and within the dose range exerting the least off-target effects [34].

There was no significant difference between hM4Di and eGFP animals in initial discrimination learning, although two rats from each virus group failed to learn (sessions to criterion: t(11) = −0.250, p = 0.807), Fig. 2A. A mixed-effects GLM was used to analyze probability correct, with drug, virus and sex as between-subject factors, session as a within-subject factor and individual rat as random factor. There was no functional consequence of ACC inhibition on mastery level performance (i.e. no significant effect of drug, virus, or interaction of drug–virus in a mixed-effects GLM; GLM formula: $\gamma \sim [1 + Drug * Session * Virus * Sex + (1 + Session | Rat)]$, Fig. 2B. A significant effect of session, interactions of session–virus and session–virus–drug were observed, Table 1. When post-hoc Bonferroni post-hoc comparisons were conducted, we found eGFP (p < 0.01) but not hM4Di (p = 0.11) groups decreased in probability across session, although VEH performance was greater than CNO in both eGFP and hM4Di groups (both p < 0.01 seconds).

In contrast to the acquisition curves that demonstrated mastery of the initial visual discrimination, all animals exhibited difficulty learning reversals, rarely achieving above 60% after 10 sessions (Fig. 2C), similar to a recent report [31]. A mixed-effects GLM was also used to analyze reversal learning. This analysis resulted in a significant drug–virus interaction ($\beta_{drug-virus} = 0.258$, t(178) = 2.56, p = 0.01). Post-hoc Bonferroni-corrected tests revealed a nonsignificant effect of drug in the hM4Di group (p = 0.07) and in the eGFP group (p = 1.0). There were other interactions with sex and session that did not result in significant effects with Bonferroni-corrected post-hoc follow-up (Table 2). However,

when we added drug order to the GLM model we found a significant drug–drug order interaction ($\beta_{drug \times order} = −0.56983$, t(162) = −3.39, p < 0.001) with an effect of drug only in the hM4Di group (post-hoc Bonferroni-corrected tests: p = 0.01 for hM4Di, p = 0.68 for eGFP). Early inhibition of ACC in the reversal phase brought performance to chance and maintained it there (Fig. 2D), in contrast to rats given VEH first which instead showed the expected tendency to follow the previous stimulus–reward assignment on early reversal sessions during VEH (i.e. more perseveration), but reached higher accuracy levels later, even when administered CNO. Interestingly, CNO had the opposite effect in eGFP animals: when administered early, rats exhibited the expected learning trajectory, but when administered CNO at 50% accuracy, performance remained at chance level (Fig. 2C). Collectively, the pattern of results reveals a differential effect of CNO in hM4Di and eGFP animals, with ACC inhibition producing a random response strategy early in reversal learning.

ACC inhibition attenuates a trial accuracy theta signal in OFC during discrimination

In a subset of learners, we collected local field potential (LFP) data in ACC and OFC (Fig. 3), acquired at 30 kHz and down sampled to 1000 Hz, and bandpass filtered for theta (5–10 Hz). We were primarily interested in measures of accuracy: correct vs. incorrect trials (Harris, Aguirre et al. [31], but we also included trial initiation and reward collection latencies in our analyses. The latter measures did not emerge as significant correlates or predictors of learning.

Based on the temporal profile of theta power changes during behavior in animals transfected with hM4Di (Fig. S2) and eGFP in ACC (Fig. S3), we baseline-subtracted theta band in ACC and OFC using the immediate pre-event period (−200 ms) leading up to the choice of the correct or incorrect stimulus. This was the moment at which the animal had already initiated a trial and

Table 1. Probability of choosing correctly during initial visual discrimination learning

	Discrimination Learning γ = probability of choosing correctly						
Formula	$\gamma \sim [1 + Drug*Session*Virus*Sex + (1 + Session \mid Rat)]$						
Coefficients	β	SE	tStat	DF	P	CI_L	CI_U
Intercept	0.75869	0.099232	7.6456	102	<0.0001	0.56186	0.95551
Virus	−0.11377	0.13422	−0.84767	102	0.3986	−0.38	0.15245
Drug	0.063004	0.11011	0.57218	102	0.56846	−0.1554	0.28141
Sex	−0.061102	0.24282	−0.25163	102	0.80183	−0.54274	0.42053
Session	−0.056029	0.011205	−5.0005	102	**<0.0001**	−0.07825	−0.03381
Virus:drug	−0.032214	0.13606	−0.23677	102	0.81331	−0.30208	0.23765
Virus:sex	0.12814	0.34095	0.37584	102	0.70782	−0.54813	0.80441
Drug:sex	−0.13438	0.23738	−0.56609	102	0.57257	−0.60523	0.33646
Virus:session	0.057054	0.016538	3.4499	102	**<0.0001**	0.024251	0.089856
Drug:session	0.04231	0.036765	1.1508	102	0.2525	−0.03061	0.11523
Sex:session	0.019448	0.12817	0.15173	102	0.8797	−0.23478	0.27368
Virus:drug:sex	−0.056662	0.35259	−0.1607	102	0.87265	−0.75603	0.64271
Virus:drug:session	−0.081474	0.039519	−2.0617	102	**0.041783**	−0.15986	−0.00309
Virus:sex:session	0.041715	0.18132	0.23006	102	0.81851	−0.31794	0.40137
Drug:sex:session	−0.029413	0.13287	−0.22138	102	0.82524	−0.29295	0.23412
Virus:drug:sex:session	0.12467	0.1982	0.62903	102	0.53074	−0.26845	0.51779

Table 2. Probability of choosing correctly during reversal learning

	Reversal Learning γ = probability of choosing correctly						
Formula	$\gamma \sim [1 + Drug*Session*Virus*Sex + (1 + Session \mid Rat)]$						
Coefficients	β	SE	tStat	DF	P	CI_L	CI_U
Intercept	0.54469	0.077476	7.0305	178	<0.0001	0.39181	0.69758
Virus	−0.20762	0.098646	−2.1047	178	**0.03672**	−0.40229	−0.01296
Drug	−0.15706	0.078148	−2.0997	178	**0.045969**	−0.31127	−0.00284
Sex	−0.31948	0.10221	−3.1258	178	**0.002072**	−0.52118	−0.11778
Session	−0.005608	0.019074	−0.29403	178	0.76908	−0.04325	0.032032
Virus:drug	0.26232	0.10267	2.5549	178	**0.011459**	0.059706	0.46493
Virus:sex	0.24254	0.13747	1.7642	178	0.079411	−0.02876	0.51383
Drug:sex	0.29399	0.10467	2.8087	178	**0.005529**	0.087436	0.50054
Virus:session	0.031054	0.020933	1.4835	178	0.1397	−0.01025	0.072363
Drug:session	0.015992	0.019692	0.8121	178	0.41782	−0.02287	0.054851
Sex:session	0.039986	0.020312	1.9686	178	0.050556	−9.79e-05	0.080069
Virus:drug:sex	−0.2572	0.1476	−1.7425	178	0.083141	−0.5485	0.034072
Virus:drug:session	−0.035597	0.023011	−1.547	178	0.12365	−0.08101	0.009812
Virus:sex:session	−0.032032	0.023057	−1.3893	178	0.16649	−0.07753	0.013468
Drug:sex:session	−0.027893	0.021945	−1.271	178	0.20537	−0.0712	0.015413
Virus:drug:sex:session	0.05759	0.028689	1.9959	178	**0.047474**	0.000645	0.11387

was fixating on the stimulus of choice. We calculated this for both discrimination and reversal phases. Despite no performance decrements following ACC inhibition in discrimination learning, normalized theta power in ACC during both correct and incorrect trials was significantly lower after CNO than after VEH in hM4Di animals (Fig. 4A). We performed 2×2 ANOVAs on this baseline-subtracted theta power comparing trial type (correct, incorrect) and drug (VEH, CNO) for discrimination sessions, separately for OFC and ACC. In ACC, we observed a significant effect of trial type [$F(1,45) = 14.237$, $p < 0.0001$] and a significant drug–trial type interaction [$F(1,45) = 5.859$, $p = 0.02$], with increases in ACC theta in both correct and incorrect trials in the VEH compared to the CNO condition (Bonferroni-corrected post-hoc comparisons, all $p < 0.035$). There was no such pattern for normalized OFC theta during discrimination, and there was a nonsignificant effect of drug [$F(1,50) = 3.965$, $p = 0.052$] (Fig. 4B). In animals with eGFP

virus, one could distinguish correct from incorrect trials from normalized theta in OFC [trial type–drug: $F(1,27) = 6.940$, $p = 0.01$, Bonferroni-corrected post-hoc comparisons, all $p < 0.01$], but not in ACC. In sum, ACC inhibition attenuated a trial accuracy theta signal in OFC during discrimination (Fig. 4B when compared to Fig. 4D), but these changes did not appreciably impact accurate performance.

Correct vs. incorrect trials are strongly differentiated by OFC, but not ACC, theta power in the reversal phase following ACC inhibition

Following ACC inhibition, correct vs. incorrect trials in reversal learning could only be distinguished by OFC, not ACC, normalized theta. As above, we performed two separate 2×2 ANOVAs, comparing trial type (correct, incorrect) and drug (VEH, CNO) for all reversal sessions, separately for OFC and ACC. In hM4Di
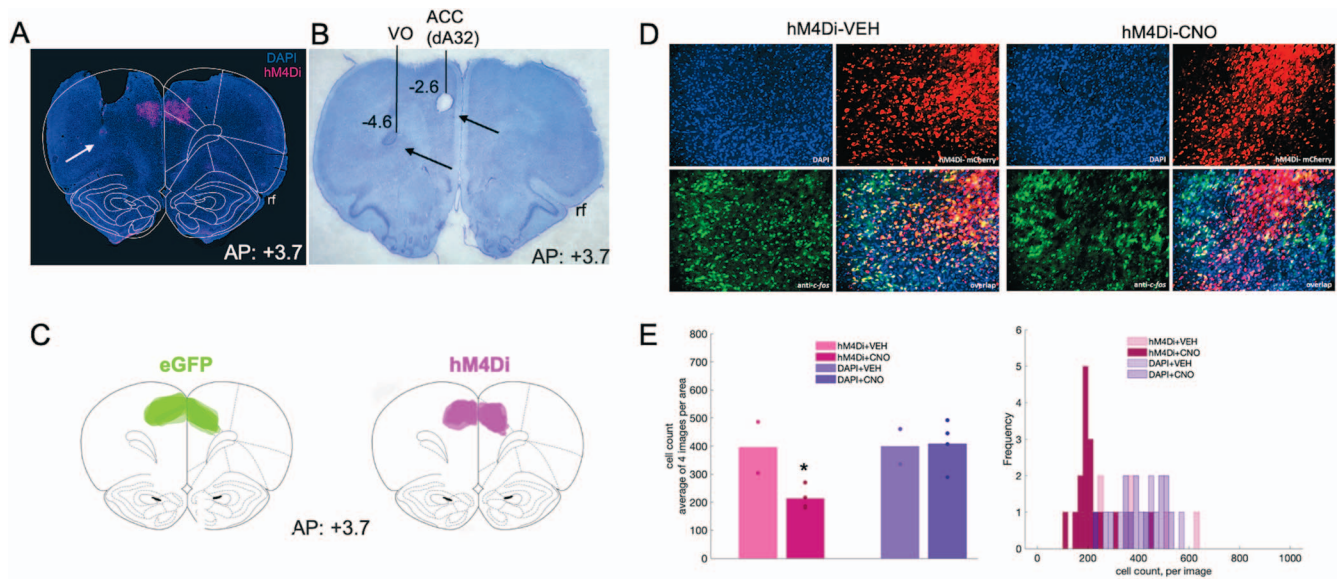
Figure 3. Inhibitory DREADDs in ACC, electrode placement in ACC and OFC, and validation of inhibition in ACC via c-fos immunohistochemistry. **(A)** Representative placement of inhibitory hM4Di DREADDs at anterior–posterior (AP) level + 3.7 relative to Bregma. Note, tissue (secondary motor cortex, M2) was removed at time of brain extraction. White arrow pointing at tip of electrode track in OFC. **(B)** Nissl-stained section showing electrolytic lesions and placement of electrode arrays targeting both ACC (dorsal area 32, dA32) and OFC (ventral orbital, VO) unilaterally in a representative animal. Black arrows pointing at tip of electrode tracks. Rf = rhinal fissure. **(C)** Reconstructions of placement of inhibitory hM4Di DREADDs (right) and eGFP null virus (left) at anterior–posterior (AP) level + 3.7 relative to Bregma. **(D)** Representative DAPI, hM4Di-mCherry, c-fos immunoreactivity and their overlap after injections of VEH (left) or CNO (right). **(E)** Mean cell count of 4 images per condition, hM4Di + VEH, hM4Di + CNO, DAPI+VEH, and DAPI+CNO. One-way ANOVA resulted in a significant decrease in c-fos–positive cells in the hM4Di-expressing areas, but not in the DAPI areas (left). Histogram of cell count frequency for each image by condition (right). *$P < 0.05$.

animals, there was a main effect of trial type on normalized OFC theta in reversal $[F(1,143) = 51.250, p < 0.0001$, all corrected Bonferroni post-hoc comparisons of correct vs. incorrect trials, all $p < 0.001]$ (Fig. 4B). This difference by trial type in normalized theta power was absent in ACC, in the hM4Di animals. In eGFP animals, there was a significant effect of trial type for normalized OFC theta $[F(1,132) = 6.21, p < 0.01]$ and ACC theta $[F(1,132) = 6.109, p = 0.01]$, with post-hoc Bonferroni-corrected comparisons revealing normalized theta was greater for correct choices compared to incorrect choices (all post-hoc tests, $p < 0.001$) (Fig. 4C, D). Thus, ACC inhibition reduced a trial accuracy theta signal during reversal learning in ACC (Fig. 4A compared to Fig. 4C), but had no impact on OFC theta. Results plotted by early vs. late reversal revealed a similar pattern (Fig. S4): there was a main effect of trial type on normalized OFC theta in reversal, whether rats received VEH first $[F(1,52) = 37.235, p = 1.33e^{-07}]$ or CNO first $[F(1,70) = 17.88, p = 6.99e^{-05}]$; all Bonferroni corrected post-hoc of correct vs. incorrect trials, $p < 0.001$. Although there was the same trend for OFC theta in eGFP animals, this difference was not statistically significant. Additionally, an analysis of individual rat (not session) changes in normalized theta revealed the same pattern of an attenuated trial accuracy theta signal during reversal learning in ACC and an intact OFC theta signal of correct vs. incorrect in the hM4Di group (Fig. S5): $[F(1,15) = 20.831, p = 6.50e^{-04}]$; all Bonferroni corrected post-hoc of correct vs. incorrect trials, $p < 0.001$.

ACC inhibition hastens incorrect choices in reversal learning

Because we were also interested in theta oscillations as they relate to speed of response, we next analyzed median latencies, i.e. to initiate a trial, to choose the correct stimulus, to choose the incorrect stimulus and to collect reward [31] in both discrimination and reversal learning. We focused on measures for which an omnibus across-phase mixed GLM analyses revealed significant

phase and drug interactions: *incorrect choice latencies* and *reward collection latencies*.

In hM4Di animals for the discrimination phase, drug, ACC theta and OFC theta were not significant predictors of incorrect choice latencies (GLM formula: $\gamma \sim [1 + Drug*ACCtheta *OFCtheta + (1 + Drug)]$). However, for the reversal phase, we found a significant effect of drug ($\beta_{drug} = -4.062$, t(67) = −2.289, $p = 0.025$; VEH slower than CNO, respective means 1.88 seconds vs 1.72 seconds, Fig. 5A) and OFC theta, not ACC theta, was a significant predictor of incorrect choice latencies (GLM: $\beta_{OFCtheta} = -89.086$, t(67) = −2.377, $p = 0.020$). For reward collection latencies in the reversal phase, we found a significant effect of drug (GLM: $\beta_{drug} = -4.217$, t(67) = −2.081, $p = 0.041$; VEH slower than CNO, respective means 0.63 seconds vs 0.52 seconds). We also observed a significant negative correlation between OFC theta power and reward collection speed under CNO (r = −0.14934, $p = 0.031$) but not under VEH conditions (r = −0.0071, $p = 0.911$). There were no other correlations between theta power and speed in hM4Di rats (Fig. 6).

In eGFP rats, we did not observe significant differences in median latencies when comparing CNO vs. VEH conditions in either the discrimination or reversal phase (Fig. 5B). Also unlike in hM4Di animals, in eGFP rats there were significant correlations between speed of correct choices (i.e. deliberation speed) and theta power in the VEH, but not CNO, conditions. During post-criterion discrimination performance, OFC theta power was negatively correlated with correct choice (r = −0.13, p = 0.02), but not under CNO conditions (r = −0.05, p = 0.331). Similarly, ACC theta power was negatively correlated with deliberation speed for correct choices under VEH (r = −0.11, p = 0.049), but not under CNO (r = −0.03, p = 0.56). In reversal learning, both ACC theta power (r = −0.16, $p = 3.55e-08$) and OFC theta power (r = −0.08, $p = 0.01$) were negatively correlated with reward collection speed under CNO, but not VEH, conditions (Fig. 7). We also found a significant drug–phase interaction (GLM: $\beta_{drug \times phase} = -3.329$, t(65) = −2.301,
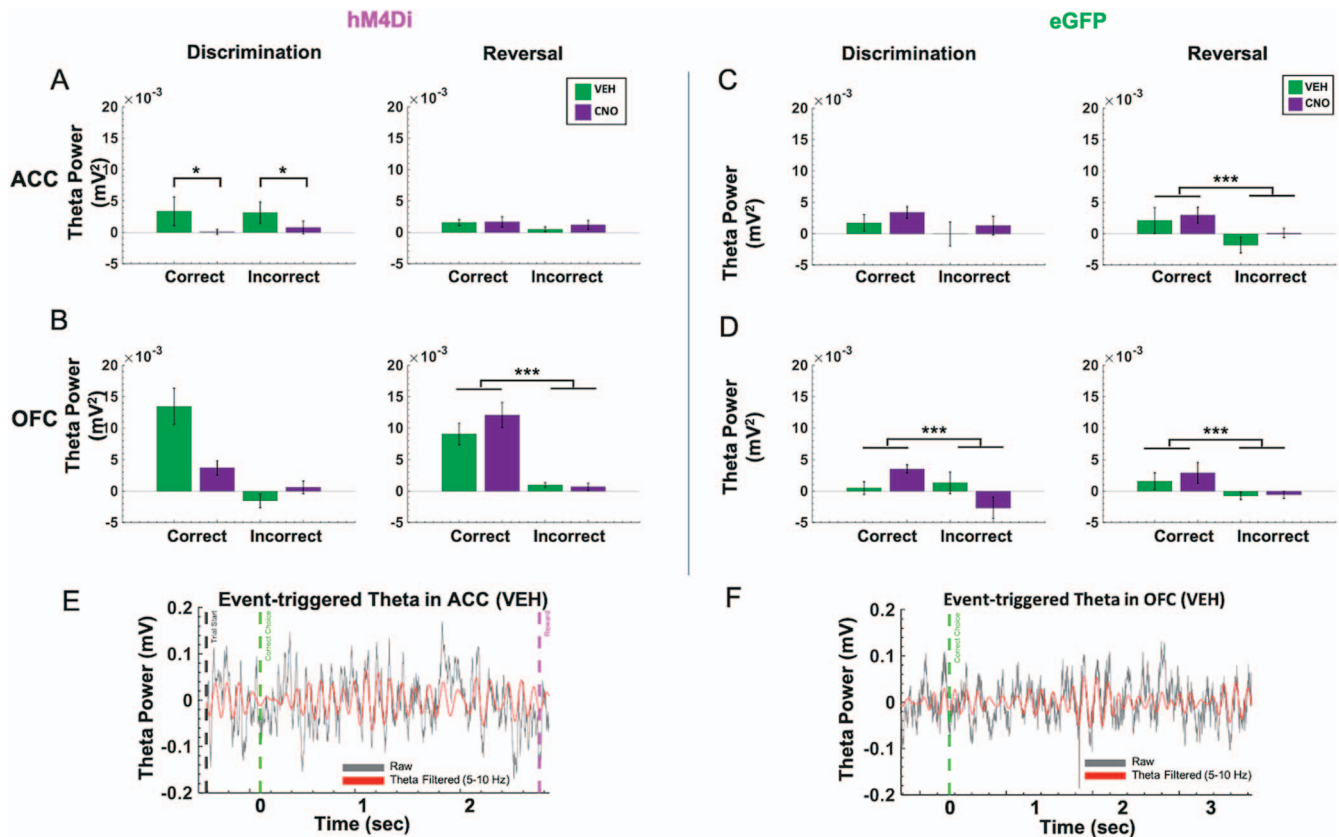
Figure 4. Trial accuracy can be distinguished only from OFC, not ACC, theta power in the reversal phase following ACC inhibition. Baseline-subtracted (normalized) theta power was calculated for correct and incorrect choices in the discrimination and reversal learning phases. (**A**) In hM4Di animals, there was no significant difference in normalized ACC theta power between correct vs. incorrect trials in the reversal phase. (**B**) in hM4Di animals, there was no significant difference in normalized theta power in OFC for correct vs. incorrect trials in discrimination, but normalized theta power in OFC did signal trial accuracy in reversal learning. (**C**) In eGFP animals, normalized theta power in ACC signaled trial accuracy in reversal learning. (**D**) In eGFP animals, normalized theta power in OFC signaled trial accuracy in both discrimination and reversal learning. (**E**) Example event-triggered raw (gray) and theta filtered (red) trace in hM4Di rat with trial start (black dotted line) and correct choice (green dotted line), and reward collection (pink dotted line) denoted for the trial. (**F**) Example event-triggered raw (gray) and theta-filtered (red) trace in eGFP rat with correct choice (green dotted line) denoted for the trial. Means ± standard error of the mean (S.E.M.) bars. ***$P < 0.001$, *$P < 0.05$ following Bonferroni-corrected post-hoc comparisons.

$p = 0.025$) for initiation latencies, however post-hoc analyses did not reveal any statistically significant differences.

Our finding speed-theta power correlations for correct choices in eGFP, and not hM4Di, animals suggest control rats are more certain of correct choices (i.e. faster correct choice, greater theta power) under VEH than under CNO during post-criterion performance, and further, that CNO enhances the relationship between reward collection speed and theta power in OFC of control rats.

Taken together, the latency analyses in learning reveal quicker deliberation speed for incorrect trials during the reversal phase following ACC inhibition. Notably, CNO had an effect on reward collection times in both eGFP and hM4Di animals, with greater theta power in OFC associated with quicker reward collection times, indicating a drug effect, not an ACC inhibition effect. However, reward collection latencies best indicate motivation levels, not deliberation speed [31]. We show here that CNO does not produce global locomotor effects which would be reflected across all trial epochs in our task, in contrast to a previous report [35].

Theta coherence during correct vs. incorrect trials

Previous studies by other groups have explored the degree to which the ACC and OFC are synchronized during effort- and value-based decisions [22, 23], with evidence suggesting significant coherence in the theta-band prior to, and after, making the correct decision. Given that we observed differences in choice

accuracy between hM4Di and eGFP animals, coherence analyses during these triggered events could provide a deeper understanding of ACC-OFC theta synchronization. Here, spectral coherence was calculated across trials after aligning the raw local field data from −2 to +2 seconds of the animal's choice, with time = 0 indicating the precise time in which the choice was made. Identical to the theta-band filtered data above, the average pre-choice period (−200 to 0 ms) was subtracted from the entire −2 to +2 second window. In the discrimination phase, a 2 × 2 (drug–virus) ANOVA revealed a significant difference between hM4Di vs. eGFP virus groups in theta coherence during correct choices in the discrimination phase [$F_{(1,39)} = 6.538$, $p = 0.01$]. However, Bonferroni-corrected post hoc t-tests did not result in a significant difference. We did not observe any statistically significant coherence differences for incorrect choices during discrimination phase (Figure S5) or for correct or incorrect choices in the reversal phase (Figure S6). Thus, we found ACC-OFC theta coherence to be fairly weak in our task despite observing significant differences in theta power during choice accuracy (Fig. 4).

DISCUSSION

In the present study, we were specifically interested in theta oscillations in frontal cortex since there have been several recent reports of involvement of oscillations in this frequency band in
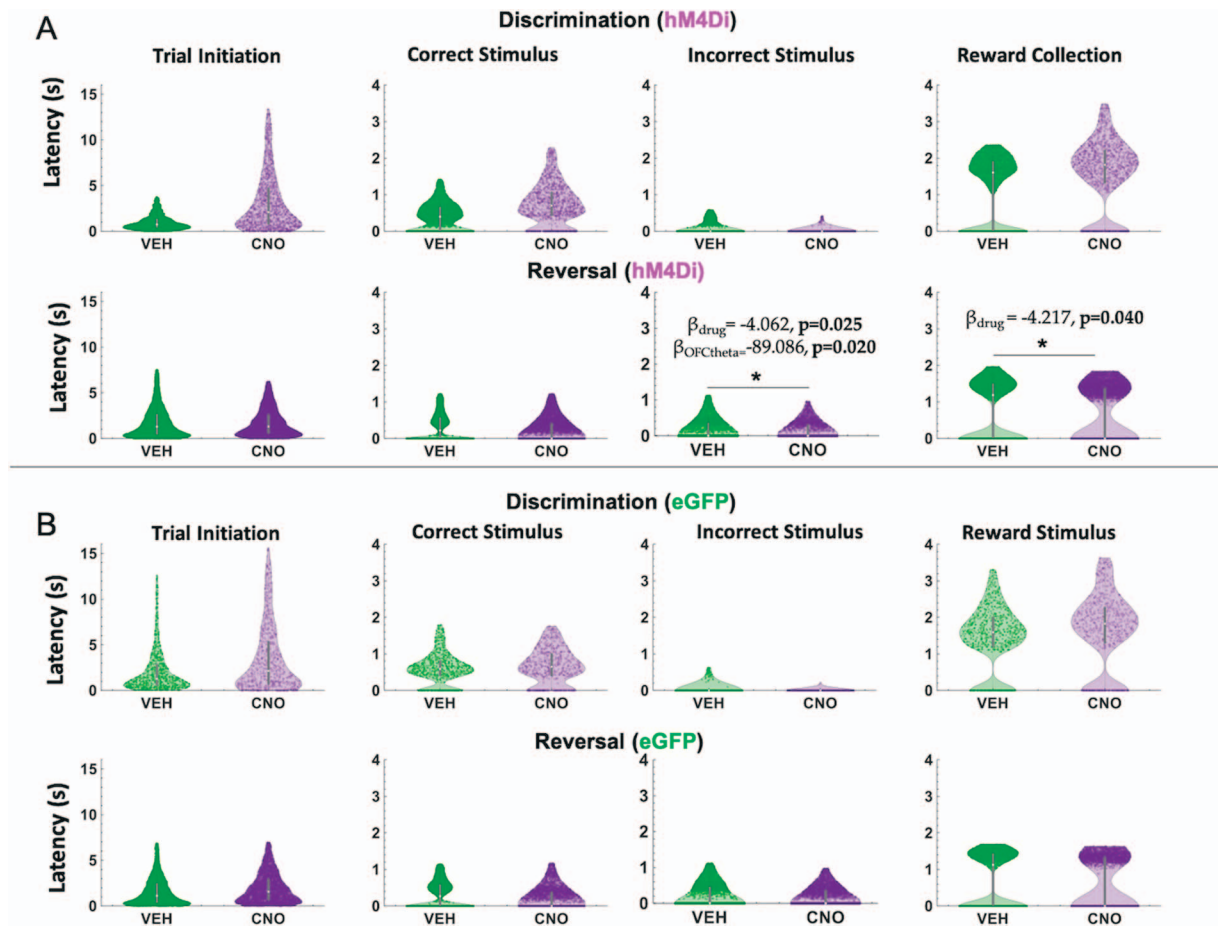
Figure 5. ACC inhibition increases speed of incorrect choices in reversal learning. Violin plots depicting individual trial median latencies in seconds (s) for each of the trial epochs in discrimination and reversal learning for hM4Di and eGFP rats. Individual dots represent trials. (**A**) In hM4Di rats, clozapine-N-oxide (CNO) administration significantly reduced the latency to choose the incorrect stimulus and collect reward in the reversal, not discrimination, phase. ACC theta did not predict speed of responses for any trial epoch in either the discrimination or reversal phase. (**B**) In eGFP rats, there was no significant hastening of incorrect choices, although there was a significant drug–phase interaction for trial initiation latencies (subsequent *post hoc* effects were not statistically significant). All significant Bbeta values are not shown, but are reported in text. *$P < 0.05$, CNO vs. VEH.

reward value and decision making in rodents [12, 21–23], and one closed-loop experiment in non-human primates demonstrating that theta oscillations in OFC are causally involved in reversal learning [26]. We found strong support for OFC theta signaling of accuracy in reversal learning. Theta is a large-scale network oscillation believed to synchronize interregional communication [36]. The precise origins of OFC theta oscillations are unknown but theta oscillations propagating from the hippocampus to the OFC have been found to encode value learning in macaques [26]. While hippocampal theta in rodents is well established to be associated with certain aspects of movement [37], others have shown that the strength of theta-band phase locking in OFC neurons follows the rat's current outcome expectation, which can be dissociated from licking responses [38]. While volume conduction from the hippocampus is a possibility, the placement of our electrodes in two areas of prefrontal cortex supports a cortical origin for the measured theta oscillations. Previous findings of OFC theta phase locking of spiking patterns also argue against simple hippocampal volume conduction [38], as well as evidence that ACC coherence with hippocampus is dynamically modulated by behavior, such that ACC theta is not just an attenuated copy of hippocampal theta [39, 40].

Here we observed a robust OFC theta power signal of accuracy in reversal learning sessions in eGFP and hM4Di animals, unper-turbed by ACC inhibition. Since LFPs in OFC, but also broadly in other areas, constitute a composite signal of multiple afferents, it is likely that another input other than ACC is more crucial to the involvement of OFC theta in reversal learning. Additionally, establishing a causal role for oscillations using chemogenetic or other viral approaches is difficult because such manipulations likely disrupt underlying firing rates, a limitation that could be bypassed by a closed-loop approach [26]. In particular, the specific silencing of CaMKII projection neurons is not a natural physiological state, and this may trigger compensatory mechanisms in downstream targets that should be investigated in high-yield recording studies in the future.

We also analyzed ACC-OFC coherence and found no evidence of significantly enhanced coherence during correct vs. incorrect choices in our task. Other groups have previously utilized a t-maze task that incorporates a spatial component to the decision process, with necessary running both prior to as well as after the choice point in the maze [22, 23]. While these authors did not determine whether running speed was correlated with ACC-OFC coherence on correct trials, it remains a possibility. High theta coherence between ACC and hippocampus has been associated with coding of an animal's current position during trajectories along a maze that may contribute to decision making [39, 41, 42]. Unlike these maze tasks, movement was much more limited
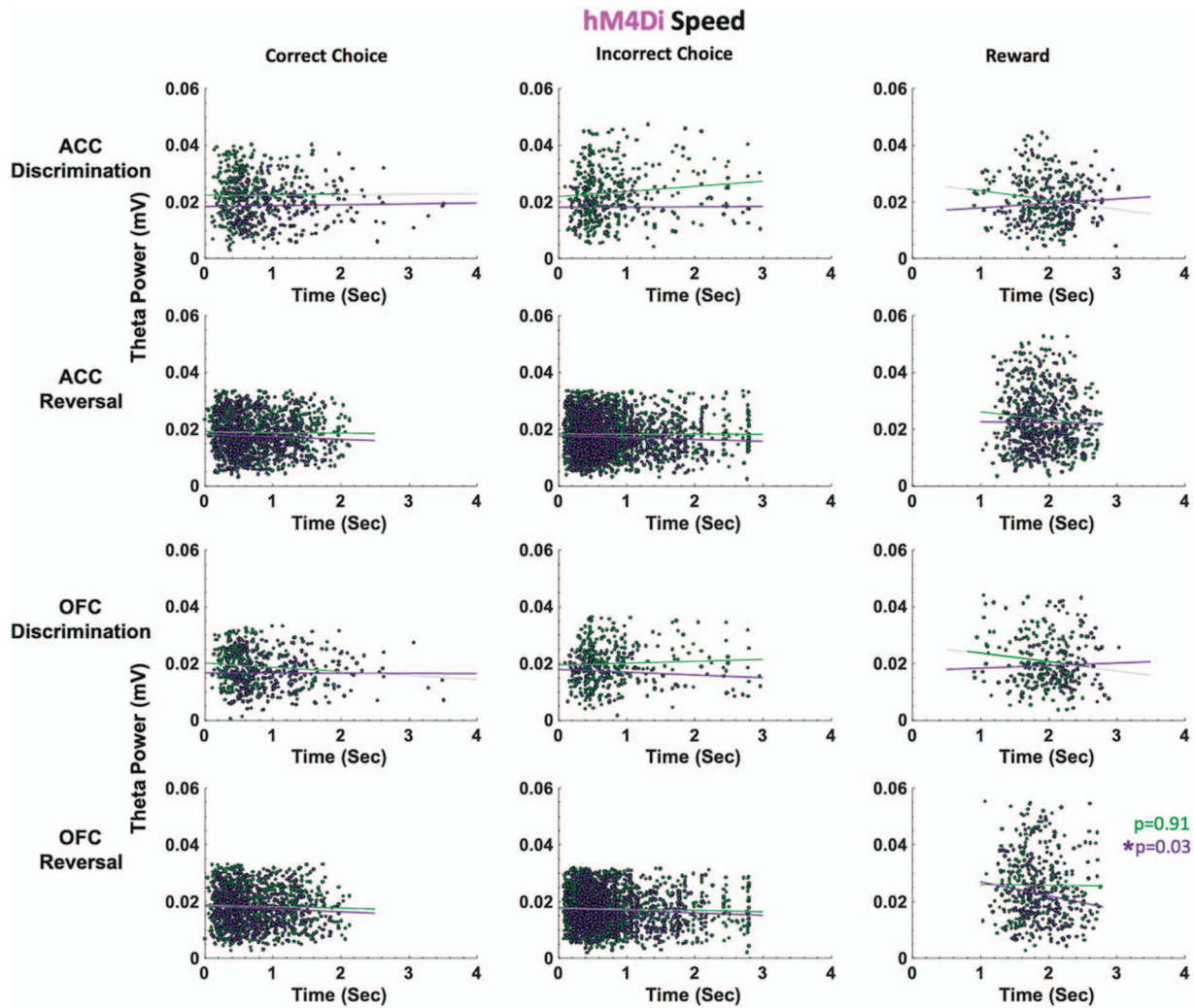
Figure 6. ACC and OFC theta power are not correlated with deliberation speed in either discrimination or reversal learning in hM4Di rats. Theta power was filtered and extracted from ACC and OFC for each event-triggered epoch (correct choice, incorrect choice, and reward port entry). Movement speed was calculated by subtracting the time at the target event (t) from the time at the previous event (t-1) in seconds (sec). OFC theta power decreased as reward collection latencies increased after CNO administration (purple) during reversal learning. This negative correlation was not observed following VEH (green). There were no significant correlations between theta power and choice of correct or incorrect stimulus. *$P < 0.05$.

in our task. Indeed, we observed the largest changes in theta power when animals made the choice of the correct stimulus, at which time they were most stationary, just having initiated a trial while fixating on the stimulus. Additionally, we found that there were significant increases in speed of incorrect choices following ACC inhibition in the same (reversal) phase when there was no theta 'read-out' of correct vs. incorrect trials. This provides some evidence that ACC inhibition eroded the distinction between correct and incorrect stimuli in the reversal phase, further elaborated below.

Various studies, mostly rodent lesion and pharmacological inactivation studies, support the idea that ACC is involved in initial cue or stimulus discrimination learning [13], but has little-to-no involvement in fully-predictive, deterministic reversal learning [14, 15]. Conversely, the prediction based on the literature is the opposite for OFC: critical for reversal, but not initial discrimination learning [11, 12]. Thus, the question about if and how ACC may be causally-involved in stimulus-based reversal learning was a question ripe for investigation. To our knowledge, we are the first to probe this using visual stimuli and with a viral-mediated approach targeting principal neurons in rats. We found that

chemogenetic inhibition of ACC indeed impacts reversal learning. The quick adoption of 50% responding is interesting and suggests that animals are either 1) newly satisfied with a 50% reward rate, or 2) unable to distinguish a discrimination vs. reversal 'task state', and thus adopt a random response strategy. The latter possibility is more likely since we found no consistent evidence of motivational changes following ACC inhibition. Labeling of the current task state and building of a cognitive map has been linked to OFC [43, 44], yet our data suggest it may also be a role of ACC, at least in rats. The latency data we report here (i.e. demonstrating that animals do not deliberate differently between correct and incorrect stimuli), support this possibility. Surprisingly, we found a drug–virus interaction in reversal learning such that CNO had a dissociable, but non-negligible effect in eGFP reversal learning. Given that this effect of CNO was specific to the late reversal phase and not observed in discrimination learning, it suggests that administration of CNO when performance is at chance may further disrupt reversal learning, given that reversal learning is modulated via a dopamine D2 mechanism [45–48].

As suggested above, theta oscillations in the hippocampus have been shown to correlate with cognitive processing, movement
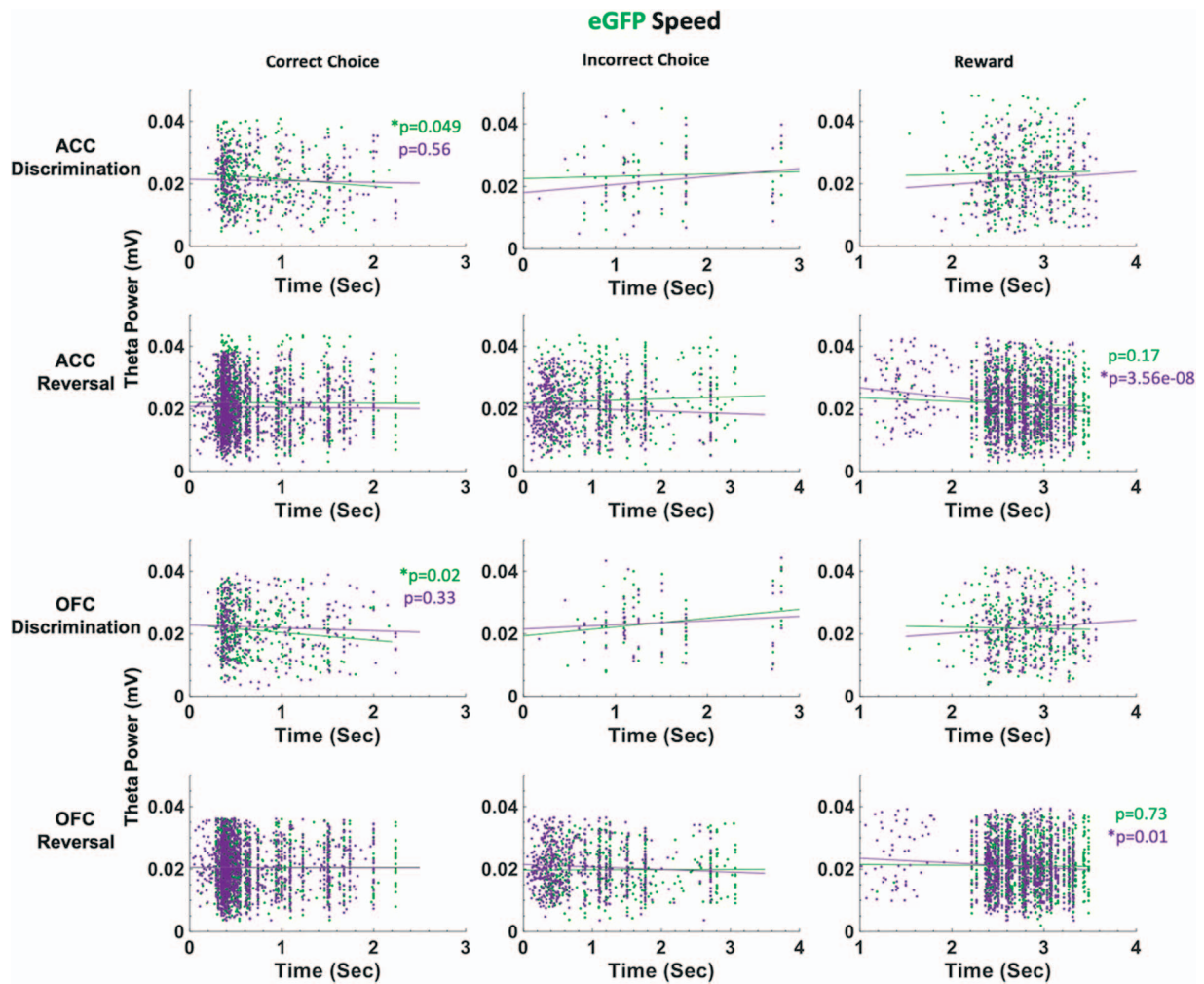
Figure 7. ACC and OFC theta power are negatively correlated with deliberation speed during discrimination performance in eGFP rats. Theta power was filtered and extracted from ACC and OFC for each event-triggered epoch (correct choice, incorrect choice, and reward port entry). Movement speed was calculated by subtracting the time at the target event (t) from the time at the previous event (t-1) in seconds (sec). OFC theta power decreased as correct choice latencies increased after VEH administration (green) during discrimination learning. This negative correlation was not observed following CNO (purple). In reversal learning, theta power in both ACC and OFC significantly decreased as reward collection latencies increased after CNO administration (purple). This negative correlation was not observed following VEH (green). There were no significant correlations between theta power and choice of correct or incorrect stimulus in reversal learning. *P < 0.05.

speed and acceleration [37, 49, 50]. Some evidence suggests theta oscillations in frontal cortex promote inter-regional coordination [51–54], yet the nature of these relationships is much less firmly established across subregions of frontal cortex. Because our rats are freely-behaving in the operant chambers, we similarly assessed speed of responding at specific trial epochs and correlations with theta power in ACC and OFC. We found that ACC inhibition hastened incorrect choices in reversal learning: this was observed only in hM4Di animals and not in eGFP animals, indicating it was indeed a result of ACC inhibition and not a broad drug effect. In a mixed-effects GLM model with ACC theta, OFC theta and drug as predictors of latencies, OFC theta and drug (CNO) emerged as the only significant predictors of these quick incorrect choices. In follow-up work, other frequency bands could be systematically probed for speed correlations since not only theta [22, 23] but also gamma [55] oscillations in frontocortical regions have been linked to preparatory motor responses to cues that predict reward. Similarly, beta increases [56] have also been observed at trial-end when working memory engagement is high (in the absence of discrete cues).

In summary, the present results suggest a role for ACC in stimulus-based reversal learning; a role similar to that historically proposed for OFC. As several psychiatric conditions manifest impairments in reversal learning including substance use disorder, obsessive compulsive disorder and schizophrenia [57–61], frontal cortex theta oscillations could be studied as biomarkers in preclinical models of these disorders.

MATERIALS AND METHODS
Animals

A total of N = 24 Long-Evans rats (Charles River Laboratories, Hollister, CA) was used for these experiments. Three animals were not included in the final dataset due to implant failure. Six male animals were surgerized to express hM4Di and after 3 weeks were euthanized shortly after being administered 3 mg/kg of clozapine-N-oxide (CNO), to assess the number of c-fos–positive cells in hM4Di-expressing areas versus neighboring (DAPI) areas following vehicle and CNO administration (see details below). Of the

n = 15 rats used for learning, eight were used for electrophysiological recordings. The latter includes male (n = 4) and female (n = 4) animals. All rats were post-natal day 40 (∼250 g) upon arrival and were pair-housed in a 12-hour reverse light/dark cycle room (lights on at 06:00 hour), maintained in 22°C to 24°C temperature conditions, with food and water available *ad libitum* prior to behavioral testing. After one week of acclimation to our vivarium, all animals were handled for 10 minutes in pairs for 5 days. After the handling period, animals were individually housed to carefully monitor food consumption under restricted access. All procedures shown in the experimental timeline (Fig. 1A) were in accordance with the recommendations of the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Chancellor's Animal Research Committee at the University of California, Los Angeles.

Viral constructs

An adeno-associated virus AAV8 driving the hM4Di-mCherry sequence under the CaMKIIa promoter was used to express DREADDs on putative projection neurons in ACC (*AAV8-CaMKIIa-hM4D(Gi)-mCherry*, packaged by Addgene, Addgene, viral prep 50477-AAV8). A virus lacking the hM4Di DREADD gene and instead containing the fluorescent tag eGFP (*AAV8-CaMKIIa-EGFP*, packaged by Addgene) was infused into ACC in four animals as a null virus control. The DREADDs-transfected and null eGFP animals underwent identical surgeries, and all of our behavior was conducted in a counterbalanced within-subject design with each animal serving as its own control (VEH compared with CNO). Collectively, this allowed us to control for non-specific effects of surgical procedures, exposure to AAV8 and non-specific effects of CNO. We note that several recent electrophysiology experiments combining neural recordings with DREADDs do not include virus (eGFP) controls [62–64], as we do here.

Behavioral apparatus

Behavioral testing was conducted in operant conditioning chambers shielded for electrophysiological recordings and outfitted with an LCD touchscreen opposing a sucrose pellet dispenser (Lafayette Instrument Co., Lafayette, IN). Sucrose pellets (45 mg; Dustless Precision Pellets F0023, Bio-Serv) were used as rewards. The inner chamber walls were modified to accommodate implanted animals by maximizing the space available (24 × 33.2 × 39.5 cm) and to improve electrophysiological signal quality. All chamber equipments were controlled by customized ABET II TOUCH software.

Pretraining

Immediately following the 5 days of handling, rats were placed on food restriction (12 g/d) to no less than 85% of their free-feeding body weight throughout the entire experiment and were weighed each day and monitored closely to not fall below this percentage. Animals were pretrained using the protocol as previously published [65]. Briefly, a series of pretraining stages: Habituation, Initiation Touch to Center (ITC) and Immediate Reward (IM), were designed to train rats to nosepoke, initiate a trial and select a stimulus to obtain a reward.

In the habituation stage, five pellets were automatically dispensed and rats were required to consume all pellets within 15 minutes to advance. In the ITC stage, the center of the touch screen displayed a white square graphic against a black background. A single sucrose pellet was dispensed with simultaneous onset of audio tone and illumination of the reward receptacle if the rat nosepoked the white stimulus, if rats took longer than

40 seconds to take this action, the white stimulus disappears without reward and trial was considered an omission followed by a 10 second inter-trial interval (ITI). Rats were required to reach a criterion of consuming 60 rewards in 45 min in order to advance to IM training. In this stage, the same white stimulus was presented as in ITC and a nosepoke indicated trial initiation. The disappearance of the white stimulus was immediately followed by the presentation of a target stimulus on the left or right side of the touchscreen (i.e. forced choice), nose poking this target stimulus was paired with a reward. An *initiation omission* was scored if rats took longer than 40 seconds to initiate the trial, a *choice omission* was scored if rats failed to nosepoke the left or right stimulus after 60 seconds. A criterion of 60 rewards consumed in 45 minutes across two consecutive days was required to advance to the discrimination phase of the behavioral task.

Surgical procedures

After completing the final pretraining stage, rats were anesthetized with isoflurane for bilateral ACC DREADDS (AAV8-CaMKIIα-hM4D(Gi)-mCherry, Addgene, Cambridge, MA, Addgene, viral prep 50477-AAV8) or eGFP (AAV8-CaMKIIa-EGFP, Addgene, Cambridge, MA, Addgene, viral prep 50469-AAV8) infusion and unilateral chronic implantation of electrode arrays in both ACC and OFC. Craniotomies were created and a 26-gauge guide cannula (PlasticsOne, Roanoke, VA) were lowered into ACC (AP: +3.7, ML: ±0.8, DV: −2.6 from skull surface), after which a 33-gauge internal cannula (PlasticsOne, Roanoke, VA) was inserted. This set of coordinates for ACC was the more anterior site of the two sites that our lab has previously utilized: more posterior ACC has been probed in effort-based tasks [32] whereas stimulus-based reversal learning has been probed with this more anterior targeting [33], constituting area 32 of cingulate cortex [27]. Animals were infused with 0.3 μL of virus at a flow rate of 0.1 μL/minute, with cannulae left in place for 5 additional minutes to allow for diffusion.

In the same surgery after DREADDs infusion, rats were implanted with a custom-designed 16-channel electrode array manufactured in-house. The array was composed of 8 stereotrodes (California Fine Wire Co., Grover Beach, CA) with four targeting the ACC and another four aimed simultaneously in the ipsilateral OFC (AP: +3.7, ML: ±2.0, DV: −4.6, from skull surface), Fig. 3A. Electrodes were lowered into the same coordinates for ACC as the viral infusions and fixed in place. Stainless steel anchor screws were placed across the skull surface and a ground/reference screw was inserted over the cerebellum. Arrays were secured in place with cyanoacrylate, bone cement (C&B Metabond, Parkell Inc., Edgewood, NY) and dental acrylic (Patterson Dental, St. Paul, MN). Animals were provided with post-operative Carprofen (5 mg/kg, *s.c.*; Zoetis, Parsippany, NJ), topical anti-biotic ointment (Water-Jel Technologies, Carlstadt, NJ) at the incision site and oral antibiotics (Sulfamethoxazole and trimethoprim oral suspension USP, Pharmaceutical Associates, Inc., Greenville, SC) daily for five days.

Drug treatment

Rats were given intraperitoneal (*i.p.*) injections of vehicle (VEH: 95% saline +5% DMSO) or CNO (3 mg/kg CNO in 95% saline +5% DMSO) 30 min prior to beginning the behavioral task [32, 33]. As shown in Fig. 1C, in the discrimination stage upon reaching 80% correct across two consecutive days without drug, rats performed four additional sessions of discrimination learning to test for mastery and successful retention. Sessions 1 and 2 after reaching 80% criterion was performed after a single i.p. CNO injection,

and Sessions 3 and 4 after a single i.p. injection of vehicle (VEH). If rats did not continue to reach criterion on Session 4 of VEH, subsequent testing sessions were administered with VEH until the criterion was exceeded after which the reversal stage would commence. Injections of VEH or CNO were administered before every reversal session. Once rats reached 50% correct across two consecutive sessions, the other drug was administered for all subsequent reversal sessions until reaching the 80% criterion. Administration of CNO or VEH for the first half (<50%) and VEH or CNO for the second half (>50%) of reversal sessions was counterbalanced across animals.

Stimulus-based discrimination and reversal learning

After 1 week of post-surgery recovery, rats began the stimulus-based discrimination learning task (Fig. 1B). This task was described in detail previously [30, 31], however, in the present design we assigned deterministic stimulus—reward assignments—correct (100%) vs incorrect (0%). Rats initiated each trial by nosepoking a white graphic stimulus in the center screen (displayed for 40 seconds). The disappearance of the initiation stimulus was immediately followed by two distinct visual stimuli ('fan' vs. 'marbles') presented on the left and right side of the touch screen (displayed for 60 seconds) pseudorandomly on each trial. One stimulus was paired with a sucrose pellet reward 100% of the time while the other stimulus was not rewarded. Trials that were not initiated within 40 seconds were scored as an *initiation omission*, whereas failure to select a choice stimulus was scored as a *choice omission*. Non-rewarded trials were followed by a 5-second time-out, and all trials including omissions were concluded with a 10-second ITI before commencing the next trial. The overall learning criterion constituted meeting the requirement of consuming 60 or more rewards along with selecting the correct option >80% of the trials within a 60-minute session across two consecutive days. As above, upon reaching criterion, rats performed an additional four sessions after receiving an injection of CNO or VEH. Rats would then advance to the reversal phase in which the previously correct stimulus was now the incorrect stimulus. Methods and criterion were identical to those described above. Rats then received injections of VEH or CNO on the first day of reversal, and on reaching 50% correct (similarly across 2 days and with 60 or more rewards collected), drug injections were switched until reaching the 80% criterion. Injections of VEH or CNO during the reversal stage were counterbalanced across animals.

Electrophysiological recordings

Local field potential (LFP) data in ACC and OFC were acquired at 30 kHz and down sampled to 1000 Hz. LFPs were bandpass filtered for theta (5–10 Hz) using the *filtfilt()* function in MATLAB. ACC DREADDs placement and electrodes placement are shown in Fig. 3. Four behavioral task epochs (e.g. trial initiation, correct choice, incorrect choice, and reward port entry) were configured as triggered event outputs from each operant chamber to the data acquisition system that were simultaneously collected with the electrophysiological recordings. Specific events were averaged over trials and over sessions, and subsequently combined across rats.

Data were acquired from a multi-channel data acquisition system (Blackrock Microsystems, LLC., Salt Lake City, UT). The digitizing headstage was connected to the electrode array, and signal was sent to the recording system through a tether and commutator (Dragonfly Inc. Ridgeley, WV). Recording sessions were conducted on all behavioral sessions preceded by an injection. Neural recordings began 15 min after drug injection followed by a 15 min baseline while the animal was inside the operant chamber. The behavioral task commenced at the end of baseline, a total of 30 min post-injection. Recording data for EGFP control animals for all trial epochs are shown in Fig. S3.

Histology

At the conclusion of the experiment, rats were anesthetized and administered electrolytic lesions at the recording sites via direct current stimulation (20 μA for 20 seconds), Fig. 3A. Rats were euthanized 3 days later by sodium pentobarbital (Euthasol, 0.8 mL, *i.p.*; Virbac, Fort Worth, TX) and brains were extracted via transcardial perfusion with phosphate buffered saline (PBS) followed by 10% buffered formalin acetate and post-fixed in this solution for 24 hours followed by 30% sucrose cryoprotection. Tissue was prepared in 40-μM thick coronal sections and either Nissl stained for verification of electrode placement (Fig. 3B) or cover slipped with DAPI mounting medium (Prolong gold, Invitrogen, Carlsbad, CA) and amplified with NMDAR1 Polyclonal Antibodies (Invitrogen for Thermo Fischer Scientific) for DREADDs verification, visualized using a BZ-X710 microscope (Keyence, Itasca, IL), and analyzed with BZ-X Viewer software. DREADDs expression (visualized by magenta fluorescence) was determined by matching histological sections to a standard rat brain atlas [66], Fig. 3C.

C-fos immunohistochemistry

Forty μm coronal sections containing ACC were first incubated overnight (16–18 hours) at 4°C in solution containing primary anti-cfos antibody (Anti-cfos (rabbit), 1:500, Abcam, Cambridge, MA, Catalog Number: ab209794), 10% normal goat serum (Abcam, Cambridge, MA, Cat. ab7481), and 0.5% Triton-X (Sigma, St. Louis, MO, Cat. T8787) in 1× PBS, followed by three 10-min washes in PBS. The tissue was then incubated for 4 hours in solution containing 1× PBS, Triton-X and a secondary antibody (Goat anti-Rabbit IgG (H + L), Alexa Fluor® 488 conjugate, 1:400, Fisher Scientific, Catalog A-11034), followed by three 10-min washes in PBS. Slides were subsequently mounted and cover-slipped with DAPI mounting medium (Prolong gold, Invitrogen, Carlsbad, CA), visualized using a BZ-X710 microscope (Keyence, Itasca, IL), and analyzed with BZ-X Viewer software. To verify DREADD-mediated inhibition of neurons in ACC we compared the number of c-fos–positive cells in the hM4Di-expressing regions to the number of c-fos–positive cells in neighboring (non-hM4Di-expressing, DAPI) areas following vehicle and CNO administration (Fig. 3DE). Four coronal sections per condition were obtained: hM4Di + VEH, hM4Di + CNO, DAPI+VEH, and DAPI+CNO. There were fewer c-fos–positive cells in hM4Di-expressing cells following CNO compared to VEH ($F(1,5) = 8.18$, $p = 0.049$), but not in DAPI-expressing cells ($F(1,5) = 0.02$, $p = 0.903$). The difference in c-fos–positive cells was not due to sampling differences: the hM4Di-expressing area was not significantly different from the non-hM4Di-expressing (DAPI) area ($t(10) = -1.536$, $p = 0.156$).

Spectral power and coherence

Local field potential (LFP) data were acquired at 30 kHz and downsampled to 1000 Hz. Traces were analyzed for signal artifacts through visual inspection and automatic artifact detection in MATLAB. LFP signals that exceeded absolute value of 1.5 mV or when summed cross-band power (2–120 Hz) exceeded the 99.98th percentile, were identified as artifact and excluded from all analyses.

Spectral power across frequency bands were determined through a fast Fourier transform via the *spectrogram()* function in MATLAB [frequency bin = 0.5 Hz, 10 seconds Hanning window]. LFP Coherence (magnitude-squared coherence) was calculated using Welch's averaged modified periodogram method with a 1-second window and a frequency resolution of 1 Hz via MATLAB function *mscohere()*. LFPs were band-pass filtered for theta (5–10 Hz) using the *filtfilt()* function in MATLAB and designed using the *designfilt()* function (stop bands: 2 Hz, 15 Hz; pass bands: 5 Hz, 10 Hz). Band-passed theta traces were synchronized from −2 to +2 seconds to each of the four behavioral task triggered events (see Data Analysis) in preparation for analyses. Once aligned, the baseline mean (−200 to 0 ms period preceding the triggered event) was subtracted from time = 0 to +2 seconds post-event for normalization. Normalized theta for accuracy trail epochs is shown in Fig. S2 and Fig. S3. ACC-OFC coherence for discrimination and reversal learning is shown in Fig. S5 and Fig. S6.

Data analysis

All behavioral and neurophysiological analyses were performed via custom-written code in MATLAB (MathWorks, Inc., Natick, MA). We first demarcated four unique task triggered events for our analyses: 1) trial initiation (center stimulus) nosepoke, 2) nosepoke to S+ (correct stimulus or action), 3) nosepoke to S- (incorrect stimulus or action), and 4) reward collection (food magazine head entry).

Learning and performance (speed/latency) data were analyzed with a series of mixed-effects General Linear Models (GLMs) (*fitglme* function; Statistics and Machine Learning Toolbox) first in omnibus analyses that included all factors and both learning phases (discrimination and reversal). For sessions to reach criterion in initial visual discrimination (when animals were not administered drug or recorded), we conducted an independent samples t-test to assess viral group differences. For theta power analyses, 2×2 ANOVAs were conducted on baseline-subtracted ACC and OFC theta power with drug (CNO, VEH)–trial type (correct, incorrect) as factors on data averaged from each session. Analyses using trial phase in this case was not treated as a within-subject variable as drug experience differed in each phase: animals received the same order of drug in discrimination (CNO first, then VEH), but drug order was counterbalanced in the reversal phase. Then, each learning phase was analyzed separately for measures where significant interactions of phase were observed. All post-hoc tests were corrected for the number of comparisons (Bonferroni). Statistical significance was noted when p-values were less than 0.05.

Major dependent variables included: percent correct (and correct and incorrect choices, probed separately) and median latencies (to initiate a trial, to make the correct choice, to make the incorrect choice, and to collect reward). Pearson correlations for latency measures and theta power were generated using the *corrcoef* function in MATLAB. Latency datapoints exceeding 2 SD were excluded. For latency values, time at trial event (t = 0) was subtracted from the previous trial event (t-1). For example, correct or incorrect choice speed was calculated by subtracting time of choice from time of trial initiation, and reward collection speed was calculated by subtracting time of reward collection from time of correct or incorrect choice, in seconds.

To confirm DREADDs efficacy, the number of c-fos–positive cells by different drug conditions (CNO, VEH) in hM4Di-expressing cells and in non-hM4Di expressing cells were analyzed with one-way ANOVAs. To ensure that c-fos–positive cells differences were not due to differences in sampling of tissue during microscopy, the area of hM4Di spread of virus was compared to that of non-hM4Di (DAPI) and analyzed using an independent samples t-test.

Author contributions

T.Y. and A.I. designed the research. T.Y., A.R. and A.I. performed the research. T.Y., J.L.R.S., C.G.A., H.T.B. and A.I. analyzed the data. T.Y., A.M.W., H.T.B. and A.I. interpreted the data. H.T.B. and A.I. acquired funding for the project. T.Y. and A.I. wrote the paper. T.Y., J.L.R.S., H.T.B., A.M.W. and A.I. edited the final version.

Conflict of Interest statement

None declared.

Supplementary Material

Supplementary data is available at *Oxford Open Neuroscience Journal* online.

References

1. Camille N, Tsuchida A, Fellows LK. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J Neurosci* 2011;**31**:15048–52
2. Luk CH, Wallis JD. Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. *J Neurosci* 2013;**33**:1864–71
3. Rudebeck PH, Behrens TE, Kennerley SW *et al.* Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* 2008;**28**:13775–85
4. Rudebeck PH, Izquierdo A. Foraging with the frontal cortex: a cross-species evaluation of reward-guided behavior. *Neuropsychopharmacology* 2021;
5. Chudasama Y, Robbins TW. Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *J Neurosci* 2003;**23**:8771–80
6. Izquierdo A, Darling C, Manos N *et al.* Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. *J Neurosci* 2013;**33**:4105–9
7. Boulougouris V, Dalley JW, Robbins TW. Effects of orbitofrontal, infralimbic and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behav Brain Res* 2007;**179**:219–28

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

9. Groman SM, Keistler C, Keip AJ *et al.* Orbitofrontal circuits control multiple reinforcement-learning processes. *Neuron* 2019;**103**:e733

10. Riceberg JS, Shapiro ML. Reward stability determines the contribution of orbitofrontal cortex to adaptive behavior. *J Neurosci* 2012;**32**:16402–9

11. Schoenbaum G, Chiba AA, Gallagher M. Neural encoding in orbitofrontal cortex and basolateral amygdala during olfactory discrimination learning. *J Neurosci* 1999;**19**:1876–84

12. Marquardt K, Sigdel R, Brigman JL. Touch-screen visual reversal learning is mediated by value encoding and signal propagation in the orbitofrontal cortex. *Neurobiol Learn Mem* 2017;**139**:179–88

13. Chudasama Y, Passetti F, Rhodes SE *et al.* Dissociable aspects of performance on the 5-choice serial reaction time task following lesions of the dorsal anterior cingulate, infralimbic and orbitofrontal cortex in the rat: differential effects on selectivity, impulsivity and compulsivity. *Behav Brain Res* 2003;**146**:105–19

14. Bussey TJ, Muir JL, Everitt BJ *et al.* Triple dissociation of anterior cingulate, posterior cingulate, and medial frontal cortices on visual discrimination tasks using a touchscreen testing procedure for the rat. *Behav Neurosci* 1997;**111**:920–36

15. Schweimer J, Hauber W. Dopamine D1 receptors in the anterior cingulate cortex regulate effort-based decision making. *Learn Mem* 2006;**13**:777–82

16. Cowen SL, Davis GA, Nitz DA. Anterior cingulate neurons in the rat map anticipated effort and reward to their associated action sequences. *J Neurophysiol* 2012;**107**:2393–407

17. Hyman JM, Holroyd CB, Seamans JK. A novel neural prediction error found in anterior cingulate cortex ensembles. *Neuron* 2017;**95**:e443

18. Mashhoori A, Hashemnia S, McNaughton BL *et al.* Rat anterior cingulate cortex recalls features of remote reward locations after disfavoured reinforcements. *elife* 2018;**7**:

19. Akam T, Rodrigues-Vaz I, Marcelo I *et al.* The anterior cingulate cortex predicts future states to mediate model-based action selection. *Neuron* 2021;**109**:e147

20. Rudebeck PH, Walton ME, Smyth AN *et al.* Separate neural pathways process different decision costs. *Nat Neurosci* 2006;**9**:1161–8

21. Amarante LM, Laubach M. Coherent theta activity in the medial and orbital frontal cortices encodes reward value. *eLife* 2021;**10**:e63372

22. Fatahi Z, Ghorbani A, Ismail Zibaii M *et al.* Neural synchronization between the anterior cingulate and orbitofrontal cortices during effort-based decision making. *Neurobiol Learn Mem* 2020;**175**:107320

23. Fatahi Z, Haghparast A, Khani A *et al.* Functional connectivity between anterior cingulate cortex and orbitofrontal cortex during value-based decision making. *Neurobiol Learn Mem* 2018;**147**:74–8

24. Hunt LT, Hayden BY. A distributed, hierarchical and recurrent framework for reward-based choice. *Nat Rev Neurosci* 2017;**18**:172–82

25. Hunt LT, Malalasekera WMN, de Berker *et al.* Triple dissociation of attention and decision computations across prefrontal cortex. *Nat Neurosci* 2018;**21**:1471–81

26. Knudsen EB, Wallis JD. Closed-loop theta stimulation in the orbitofrontal cortex prevents reward-based learning. *Neuron* 2020;**106**:e534

27. van Heukelum S, Mars RB, Guthrie M *et al.* Where is cingulate cortex? A cross-species view. *Trends Neurosci* 2020;**43**:285–99

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

29. Hoover WB, Vertes RP. Projections of the medial orbital and ventral orbital cortex in the rat. *J Comp Neurol* 2011;**519**:3766–801

30. Aguirre CG, Stolyarova A, Das K *et al.* Sex-dependent effects of chronic intermittent voluntary alcohol consumption on attentional, not motivational, measures during probabilistic learning and reversal. *PLoS One* 2020;**15**:e0234729

31. Harris C, Aguirre CG, Kolli S *et al.* Unique features of stimulus-based probabilistic reversal learning. *Behav Neurosci* 2021;**135**:550–70

32. Hart EE, Blair GJ, O'Dell TJ *et al.* Chemogenetic modulation and single-photon calcium imaging in anterior cingulate cortex reveal a mechanism for effort-based decisions. *J Neurosci* 2020;**40**:5628–43

33. Stolyarova A, Rakhshan M, Hart EE *et al.* Contributions of anterior cingulate cortex and basolateral amygdala to decision confidence and learning under uncertainty. *Nat Commun* 2019;**10**:4704

34. Jendryka M, Palchaudhuri M, Ursu D *et al.* Pharmacokinetic and pharmacodynamic actions of clozapine-N-oxide, clozapine, and compound 21 in DREADD-based chemogenetics in mice. *Sci Rep* 2019;**9**:4522

35. MacLaren DA, Browne RW, Shaw JK *et al.* Clozapine N-oxide administration produces behavioral effects in long-Evans rats: implications for designing DREADD experiments. *eNeuro* 2016;**3**:ENEURO.0219–16.2016

36. Buzsaki G. Theta rhythm of navigation: link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 2005;**15**:827–40

37. Kennedy JP, Zhou Y, Qin Y *et al.* A direct comparison of theta power and frequency to speed and acceleration. *J Neurosci* 2022;**42**:4326–41

38. van Wingerden, Vinck M, Lankelma J *et al.* Theta-band phase locking of orbitofrontal neurons during reward expectancy. *J Neurosci* 2010;**30**:7078–87

39. Remondes M, Wilson MA. Cingulate-hippocampus coherence and trajectory coding in a sequential choice task. *Neuron* 2013;**80**:1277–89

40. Young CK, McNaughton N. Coupling of theta oscillations between anterior and posterior midline cortex and with the hippocampus in freely behaving rats. *Cereb Cortex* 2009;**19**:24–40

41. Remondes M, Wilson MA. Cingulate-hippocampus coherence and trajectory coding in a sequential choice task. *Neuron* 2014;**81**:1214

42. Zielinski MC, Shin JD, Jadhav SP. Coherent coding of spatial position mediated by theta oscillations in the hippocampus and prefrontal cortex. *J Neurosci* 2019;**39**:4550–65

43. Costa KM, Scholz R, Lloyd K *et al.* The role of the lateral orbitofrontal cortex in creating cognitive maps. *Nat Neurosci* 2022;

44. Wilson RC, Takahashi YK, Schoenbaum G *et al.* Orbitofrontal cortex as a cognitive map of task space. *Neuron* 2014;**81**:267–79

45. Baldessarini RJ, Centorrino F, Flood JG *et al.* Tissue concentrations of clozapine and its metabolites in the rat. *Neuropsychopharmacology* 1993;**9**:117–24

46. Lee B, Groman S, London ED *et al*. Dopamine D2/D3 receptors play a specific role in the reversal of a learned visual discrimination in monkeys. *Neuropsychopharmacology* 2007;**32**:2125–34

47. Linden J, James AS, McDaniel C *et al*. Dopamine D2 receptors in dopaminergic neurons modulate performance in a reversal learning task in mice. *eNeuro* 2018;**5**:ENEURO.0229–17.2018

48. Manvich DF, Webster KA, Foster SL *et al*. The DREADD agonist clozapine N-oxide (CNO) is reverse-metabolized to clozapine and produces clozapine-like interoceptive stimulus effects in rats and mice. *Sci Rep* 2018;**8**:3840

49. O'Keefe J, Nadel L *The hippocampus as a cognitive map*. New York, Clarendon Press: Oxford University Press, 1978

50. Vanderwolf CH. Hippocampal electrical activity and voluntary movement in the rat. *Electroencephalogr Clin Neurophysiol* 1969;**26**: 407–18

51. Cavanagh JF, Frank MJ. Frontal theta as a mechanism for cognitive control. *Trends Cogn Sci* 2014;**18**:414–21

52. Jones MW, Wilson MA. Theta rhythms coordinate hippocampal-prefrontal interactions in a spatial memory task. *PLoS Biol* 2005;**3**:e402

53. Rajan A, Siegel SN, Liu Y *et al*. Theta oscillations index frontal decision-making and mediate reciprocal frontal-parietal interactions in willed attention. *Cereb Cortex* 2019;**29**:2832–43

54. Siapas AG, Lubenov EV, Wilson MA. Prefrontal phase locking to hippocampal theta oscillations. *Neuron* 2005;**46**:141–51

55. Donnelly NA, Holtzman T, Rich PD *et al*. Oscillatory activity in the medial prefrontal cortex and nucleus accumbens correlates with impulsivity and reward outcome. *PLoS One* 2014; **9**:e111300

56. Schmidt R, Herrojo Ruiz M, Kilavik BE *et al*. Beta oscillations in working memory, executive control of movement and thought, and sensorimotor function. *J Neurosci* 2019b;**39**:8231–8

57. Brigman JL, Ihne J, Saksida LM *et al*. Effects of subchronic phencyclidine (PCP) treatment on social behaviors, and operant discrimination and reversal learning in C57BL/6J mice. *Front Behav Neurosci* 2009;**3**:2

58. Izquierdo A, Jentsch JD. Reversal learning as a measure of impulsive and compulsive behavior in addictions. *Psychopharmacology* 2012;**219**:607–20

59. Leeson VC, Robbins TW, Matheson E *et al*. Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: stability over six years and specific associations with medication type and disorganization syndrome. *Biol Psychiatry* 2009;**66**: 586–93

60. Remijnse PL, Nielen MM, van Balkom *et al*. Reduced orbitofrontal-striatal activity on a reversal learning task in obsessive-compulsive disorder. *Arch Gen Psychiatry* 2006;**63**: 1225–36

61. Winstanley CA, Olausson P, Taylor JR *et al*. Insight into the relationship between impulsivity and substance abuse from studies using animal models. *Alcohol Clin Exp Res* 2010;**34**:1306–18

62. Alexander GM, Brown LY, Farris S *et al*. CA2 neuronal activity controls hippocampal low gamma and ripple oscillations. *eLife* 2018;**7**:

63. Liu Y, McAfee SS, Heck DH. Hippocampal sharp-wave ripples in awake mice are entrained by respiration. *Sci Rep* 2017;**7**:8950

64. Schmidt B, Duin AA, Redish AD. Disrupting the medial prefrontal cortex alters hippocampal sequences during deliberative decision making. *J Neurophysiol* 2019a;**121**:1981–2000

65. Stolyarova A, Izquierdo A. Complementary contributions of basolateral amygdala and orbitofrontal cortex to value learning under uncertainty. *elife* 2017;**6**:

66. Paxinos G, Watson C *The rat brain in stereotaxic coordinates*. Amsterdam; Boston: Academic Press/Elsevier, 2007