Published in final edited form as:

*Curr Biol*. 2022 November 07; 32(21): 4675–4687.e5. doi:10.1016/j.cub.2022.09.022.

# Orbitofrontal cortex populations are differentially recruited to support actions

**Christian Cazares**[1], **Drew C. Schreiner**[2], **Mariela Lopez Valencia**[2], **Christina M. Gremel**[1,2,*]

[1]The Neurosciences Graduate Program, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA

[2]Department of Psychology, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA

## Summary

The ability to use information from one's prior actions is necessary for decision-making. While Orbitofrontal cortex (OFC) has been hypothesized as key for inferences made using cue and value-related information, whether OFC populations contribute to the use of information from volitional actions to guide behavior is not clear. Here, we used a self-paced lever-press hold down task in which mice infer prior lever press durations to guide subsequent action performance. We show that activity of genetically identified lateral OFC subpopulations differentially instantiate current and prior action information during ongoing action execution. Transient state-dependent lOFC circuit disruptions of specified subpopulations reduced the encoding of ongoing press durations but did not disrupt the use of prior action information to guide future action performance. In contrast, a chronic functional loss of lOFC circuit activity resulted in increased reliance on recently executed lever press durations and impaired contingency reversal, suggesting the recruitment of compensatory mechanisms that resulted in repetitive action control. Our results identify a novel role for lOFC in the integration of action information to guide adaptive behavior.

## eTOC

Cazares et al. show that lateral OFC is necessary for action-related information used in decision-making. Actions are differentially represented by activity of distinct lOFC populations, with activity perturbations preventing the encoding, but not use, of prior action information. lOFC lesions recruit compensatory repetitive action control.

*Corresponding Author: Christina M. Gremel, Ph.D., University of California San Diego, 9500 Gilman Drive, Mail Code 0109, La Jolla, CA 92093-0109, 858-594-8933, cgremel@ucsd.edu.

Author Contributions

C.C.: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Visualization, Writing - original draft, Writing - review and edition. D.C.S.: Formal analysis and Writing - review and editing. M.L.V.: Investigation and Writing - review and editing. C.M.G.: Conceptualization, Methodology, Supervision, Funding acquisition, Project administration, Visualization, Writing - original draft, Writing - review and editing.

Declaration of interests

The authors declare no competing interests.

Inclusion and Diversity

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in their field of research or within their geographical location. One or more of the authors of this paper received support from a program designed to increase minority representation in their field of research.

**Keywords**

action; orbitofrontal cortex; experience; parvalbumin; fiber photometry

## Introduction

Flexible decision-making requires successful use of information derived from past experiences.[1–4] Orbitofrontal cortex (OFC) has been hypothesized to process inferred information relevant to ongoing task demands, integrating inferences into a "cognitive map" to support ongoing decision-making processes.[5–11] Past investigations have supported this hypothesis, showing OFC activity contributes to inferred information derived from externally-derived sources, such as with predictive cues that can elicit behavior,[12–16] cued choices,[17–22] and outcome value.[23–30] However, prior actions can also be used as information for inferences critical to adaptive control[3,31–33] and whether OFC populations are recruited for use of such action-related information has been debated.[34] This is important to resolve as OFC activity is disrupted across psychiatric disorders involving aberrant action control.

Volitional actions provide one the ability to dictate opportunities and achieve desired outcomes.[35] This differs from situations where behavior can be elicited or signaled by external sources in the environment. Instead, self-generated actions are organized and initiated based on inferences that arise within one's self.[36] For example, a road sign can signal which way to walk to get ice cream, or one can infer from prior experiences which direction to go. Depending on recent experiences, one can repeat actions to exploit a known rule, or modify an action to explore for new rules,[37,38] allowing one to adjust behavior from one decision to the next. However, whether such action-related inferences recruit OFC-based contributions is less clear. On one hand, prior investigations have observed modulation of lateral OFC (lOFC) neurons during actions[39–42] and found that disrupting lOFC activity perturbs actions sensitive to decreases in expected outcome value.[40,43–46] Additional work has suggested lOFC is recruited when action-outcome contingencies change during learning.[47] In contrast, otherl studies have suggested that OFC populations may not participate in action processes per se, but instead are only recruited when cue-related, outcome-related, or value-related information is also present.[48,13,49–55] In support of the latter hypothesis, broad (i.e. not population specific) chemical-induced lOFC inactivation in marmosets was found to enhance choice sensitivity to changes in action-outcome contingencies, raising the possibility that OFC performs functions that compete with action control processes.[56] However, procedures often used to assess or degrade action-outcome contingencies or to decrease the value of associated outcomes do not provide a way to examine adjustments to the action itself, independently from adjustments to the relationship between an action and its associated outcome. Furthermore, relatively longer-term (i.e. lesions or whole-session manipulations) and non-specific lOFC activity disruptions found in these aforementioned studies may have facilitated compensatory mechanisms that could assume responsibility for the observed behavioral disparities.[57,44,58–60] Thus, the specific contributions of lOFC to action-related information, if any, remain ambiguous.

Lateral OFC is widely innervated by cortical, thalamic, and subcortical areas,[61–66] with incoming afferents synapsing onto various cortical cell types that include excitatory projection neurons and local interneuron populations that shape local network rhythmicity and neuronal firing.[67–71] Despite this vast interconnectivity, little is known about how information used for inferences is integrated within these lOFC microcircuits. As different cell-types may receive similar inputs, and thus potentially similar information, genetically distinct lOFC subpopulations could show functional homogeneity or differential representation of information used for inferences that guide adaptive behavior. Here we investigated whether lOFC projection and local inhibitory populations are important for volitional action control. We used a self-paced instrumental task that allowed us to examine adjustments to action control while keeping broad action-outcome relationships stable.[33,42] Our data suggests lOFC performs computations that contribute to the encoding of action history to guide adaptive behaviors.

## Results

### Mice learned to adjust self-generated lever presses using inferred action-related experience

Behavior is shaped in real time by its history. We adapted a lever-press hold down task in which mice had to learn to hold down a lever press for an arbitrary, unsignaled and predetermined minimum amount of time (Figure 1A). Importantly, reward delivery only occurred immediately after the termination of any lever press that exceeded this duration criterion and there were no cues predictive of reward.[33,42,72–75] Lever pressing was self-initiated, self-paced, thus mice had only their experience, including prior lever press durations, to guide subsequent action performance.

Lever press duration criterion was set at >800 ms for five daily sessions, followed by five daily sessions with a >1600 ms duration criterion (Figure 1B). A representative session from a well-trained mouse during a 1600 ms criteria day shows variability in the duration and frequency of lever presses made across the session (Figure 1C). Mice reduced the number of total lever presses made (Figure 1D, one-way RM ANOVAs for 800 ms and 1600 ms training durations, Fs > 35.84, ps < 0.0001) and decreased response rates across each duration criteria (Figure 1E, one-way RM ANOVAs Fs > 18.16, ps < 0.0001). Mice increased successful task performance within each duration criteria rule, as shown by an increase in the percentage of total presses made that exceeded the minimum duration criterion (Figure 1F; one-way RM ANOVAs for 800 ms and 1600 ms training durations Fs > 31.32, ps < 0.0001). In addition, we observed rightward shifts in the distributions towards longer press durations when duration criteria shifted (Figure 1G; two-way RM ANOVA, main effect of Duration Bin $F_{1.943, 69.96} = 336.5$, $p < 0.0001$, main effect of Criterion $F_{1.067,38.40} = 7.061$, $p < 0.05$, and an interaction (Duration Bin * Criterion) $F_{2.622, 94.38} = 104.8$, $p < 0.0001$). Outcome devaluation testing showed initiation of lever-pressing was goal-directed (Figures 1H, Paired t-test, $t_{16} = 2.482$, $p < 0.05$, and S1A–B). Of note, the percentage of successful lever presses did not differ between valuation states (Figure S1C), suggesting different behavioral mechanisms may control remaining lever presses

made.[33,42] Thus, broad behavioral performance measures suggested that mice used inferred contingency and expected outcome information to guide their lever press behavior.

However, it was unclear what information mice were using to adjust lever press performance. As previously reported,[33] the behavior of mice violated the scalar property of timing (ratio of median and interquartile range (IQR) of lever press durations) (Figure 1I; Paired t-test, $t_{36} = 2.588$, $p < 0.05$). This suggested that mice may not have exclusively timed each lever press independently, but rather that lever press durations were also influenced by preceding lever press durations as well as other sources of information derived from recent experience, including prior reward delivery, prior checking behavior, as well as the time passed between lever presses (interpress-interval) (Figure 1J). We built linear mixed effect models (LMEs) that measured the predictive relationship of these behavioral events on the subsequent lever press duration (n) and compared LME regression coefficients (β) against lever press order-shuffled data via permutation testing.

We found that mice relied on prior experiential information to guide lever pressing. First, sequential lever press durations (n − 1) were related to one another (Figure 1K; $p < 0.01$, Table S1, top), with the predictive relationship decaying up to the 10th prior (n − 10) lever press duration (Figure S1D; ps < 0.05, Table S1, middle), confirming mice inferred prior lever press durations to adjust future responding. Mice made shorter presses after reward delivery (n − 1 Outcome), potentially indicative of titrating lever press durations for performance success (Figure 1L; $p < 0.01$).[33,74] Checking behavior, indexed via a head entry into the food receptacle, increased the subsequent lever press duration (Figure 1M; $p < 0.01$). Furthermore, the longer the interval in between presses, the shorter the subsequent lever press duration (Figure 1N; $p < 0.01$). Importantly, and in line with a prior report,[33] we found that the relationship between sequential presses (β coefficient for n and n−1) was modified by whether the animal made a head entry (i.e. checking behavior) as well as how much time had elapsed (i.e. interpress interval) between presses (Table S1, bottom). In contrast, reward delivery did not alter the relationship between current (n) and prior (n − 1) lever press durations (Table S1, bottom), suggesting that the presence or absence of reward did not change how mice used prior lever press duration information to guide subsequent performance. The use of this experiential information improved performance. We tested LME model performance using individual session data and found a positive relationship between model $R^2$ for an individual and that individual's overall session performance efficiency (Figures S1E and S1F). Subjects used experiential information to adjust behavior even during early learning; for example mice relied on prior duration information to a similar degree between early (i.e. first few 800ms days) and late acquisition (i.e. last few 1600 ms days) (Figure S1G). However, the influence of other sources of experiential information, such as prior outcome, checking behavior, and interpress interval, increased as training continued (Figures S1H–S1J). The above findings replicate previous results showing that when behavior is largely uninstructed, mice rely on numerous sources of experiential information to guide volitional action control,[33] including inferences about prior action performance.

### lOFC populations differentially encode actions and action-related information

We next examined whether OFC reflects experiential information during decision-making, particularly lever press duration information. To avoid head fixation effects on context-dependent behaviors,[76] we monitored OFC population $Ca^{2+}$ activity of Calcium/calmodulin-dependent protein kinase II (CaMKII+) projection populations (rAAV5/PAAV-CaMKIIa-GCaMP6s) using *in vivo* fiber photometry in freely-moving mice as they lever pressed during 1600 ms duration criterion training (Figures 2A and S2). A perievent histogram of $Ca^{2+}$traces ordered by press duration revealed that CaMKII+ OFC projection population activity was modulated at select epochs relative to lever press initiation and execution (Figure 2B). We segmented CaMKII+ fluorescence activity traces by whether or not the lever press was eventually rewarded. Group averaged CaMKII+ OFC projection population activity was modulated prior to the onset of a lever press, similarly to previously reported single-unit recordings.[40,42] Permutation testing[77] revealed that pre onset lOFC $Ca^{2+}$ activity differed with respect to its eventual outcome (Figure 2C; ps < 0.05). Future success-related differences persisted during the lever press (Figure 2D; ps < 0.05) and success-related differences were observed after the lever press release (Figure 2E; ps < 0.05). Indeed, we observed greater levels of CaMKII+ OFC projection population activity in head entries that followed a successful lever press than in head entries following an unsuccessful lever press, corresponding to a time point during which mice had access to food pellet-related sensory and consummatory information (Figure 2F; ps < 0.05).

To investigate whether information about eventual performance outcomes originated from aspects of experiential information, we built LME models which aimed to predict lever press aligned changes in calcium activity given the current lever press duration and prior experiential information, with a focus on action-related information (Figure 2G). Prior to lever press onset, we found a significant relationship between CaMKII+ OFC projection population activity and the upcoming (n) lever press duration (Figure 2H). This significant relationship was also found while the animals held down the lever (Figure 2I) and was still present at termination of the lever press (Figure 2J). In other words, prior to lever press onset, greater increases in lOFC CaMKII+ activity were associated with longer durations of the upcoming action. However, during lever press execution, lower levels of activity corresponded to longer lever presses. At lever press offset, longer lever presses were associated with increased lOFC CamKII activity. In contrast, lOFC CamKII+ activity did not appear to maintain prior duration information, as there was no relationship between the prior lever press duration and current lOFC CamKII+ calcium activity at lever press onset (Figure 2H). However, there were significant relationships between current lOFC CamKII+ calcium activity and prior (n − 1) lever press during lever press execution (Figure 2I) as well as at lever press offset (Figure 2J). We also found that CaMKII+ OFC projection population activity during these lever press epochs was modulated by whether the prior lever press was rewarded or not, whether a checking head entry was made, as well as the time from prior lever press (Table S2). Together, our results suggest that current and prior action-related information as measured by lever press durations, as well as broader experiential information, is differentially encoded by CamKII+ projection populations in the OFC.

OFC projection circuits do not act in isolation. Local GABAergic interneurons are crucial for the control of local circuit inhibition[68,70,71] with Parvalbumin (PV+) interneurons being critical for spike timing and synchronizing network oscillations.[67,69] As cortical projection and local inhibitory populations receive long-range cortical input,[64,65] it may be that PV+ lOFC inhibitory population activity reflects experiential information similar to that of CaMKII+ lOFC projection populations. We performed fiber photometry experiments monitoring $Ca^{2+}$ activity of virally targeted PV+ lOFC interneuron populations (rAAV5/pAAV.CAG.Flex.GCaMP6s.WPRE.SV40) in freely moving $PV^{cre}$ mice as they performed the lever-press hold down task (Figure 2K). A peri-event histogram of baseline normalized traces ordered by press duration suggested that PV+ lOFC neuron population activity patterns, notably differed from CaMKII+ lOFC projection populations (Figure 2L). Permutation testing revealed that group averaged lOFC PV+ $Ca^{2+}$ activity was modulated prior to the onset of a lever press, with smaller reductions of activity observed with lever presses that would be rewarded ($ps < 0.05$) (Figure 2M). Reward-related differences in lOFC PV+ $Ca^{2+}$ activity persisted during ongoing lever press execution (Figure 2N) and after lever press offset (Figure 2O), increasing to a larger degree for rewarded than unrewarded press durations ($ps < 0.05$). In contrast to CaMKII+ lOFC projection populations, lOFC PV+ $Ca^{2+}$ activity associated with head entries made following lever press release reached similar levels regardless of the presence of a reward (Figure 2P).

We again built LME models which aimed to predict lever press aligned changes in PV+ $Ca^{2+}$ activity given prior and current lever press durations, as well as other sources of experiential information (Figure 2Q). In contrast to the CaMKII+ lOFC projection population, PV+ lOFC interneuron population $Ca^{2+}$ activity prior to (Figure 2R) and throughout the lever press (Figure 2S) was not predictive of ongoing (n) or prior (n − 1) lever press durations, prior checking behavior, nor the inter-press interval, but was predictive of whether the prior press was rewarded or not (Table S3). However, at lever press offset, PV+ lOFC interneuron population activity was predictive of the duration of the lever press that was just executed (n), as well as prior checking behavior and lever press outcome (Figure 2T and Table S3). Our results suggest that, unlike CamKII+ projection populations, PV+ inhibitory population activity in lOFC largely reflects outcome-related information during lever pressing and consequence-related information after lever press termination. Together, our results suggest that inferred action information is differentially encoded by lOFC sub-populations.

## lOFC encodes action-related information to modify behavior

To test whether lOFC activity functionally contributes to actions, we selectively disrupted local OFC activity in a temporally-specific manner during lever press execution. We used an optogenetic approach to bilaterally activate PV+ lOFC inhibitory populations with an excitatory opsin (rAAV5/Ef1a-DIO-hChR2(H134R)-eYFP) to inhibit local lOFC projection population activity (Figure 3A).[29,78] Optical stimulation was behaviorally-dependent on the execution of a lever press, such that the initiation of every lever press, independent of their eventual duration, triggered light delivery (470 nm 20 Hz, 5 ms pulses) that continued until press termination (Figure 3B). Stimulation days occurred after task acquisition (Figure 3C). In days in which light was delivered, ChR2 mice maintained similar rates of responding

(Figure 3D; $p > 0.05$), but reduced the percentage of rewarded lever presses compared to fluorophore controls (Figure 3E; two-way RM ANOVA, main effect of Treatment only $F_{1, 12} = 4.990$, $p < 0.05$). Comparisons of lever press duration distributions suggested that light activation altered the distribution pattern of lever press durations in ChR2 mice (Figure 3F; two-way RM ANOVA, main effect of Duration Bin $F_{1.813, 21.76} = 56.73$, $p = 0.0001$; marginally significant interaction (Duration Bin * Treatment) $F_{9, 108} = 1.967$, $p = 0.05$). The above data suggest disruption of lOFC activity during action execution impaired successful performance.

We next examined whether this disruption of OFC activity impaired performance in part by changing the predictive relationship between prior (n − 1) and ongoing (n) lever press durations. We added a term to our behavioral LME model that accounted for the presence or absence of the excitatory opsin (Treatment) in each animal. We found a significant interaction between the prior lever press duration and the presence of the opsin ($Duration_{n-1}$ * Treatment) in predicting subsequent lever press durations (Table S4; $p < 0.05$). A representation of group-segmented β coefficients showed a reduced relationship between prior (n − 1) and current (n) lever press durations in ChR2 animals compared to fluorophore controls (significant compared to 1000 group-shuffled data) (Figure 3G; $p < 0.01$). The above suggests lOFC activity supports action-related information.

The above approach prevented us from determining whether lOFC activity contributed to the encoding of action information, or to the use of prior action information to guide ongoing performance. Therefore, we next directly examined the acute behavioral contributions of lOFC activity patterns and their support for the encoding and/or use of action-related information. We bilaterally expressed an excitatory opsin in CamKII+ projection neurons (rAAV5/CamKII-hChR2(H134R)-eYFP-WPRE) to induce non-physiological increases in lOFC activity during lever pressing (i.e., when activity is normally decreased) in a small subset of actions made during the session by pairing light activation (470 nm 20 Hz, 5 ms pulses) only to every 7th lever press (Figure 4A). This allowed us to examine whether light activation during prior (n − 1) press would prevent its encoding and therefore reduce its contribution to the ongoing (n) duration, and/or would light activation during the ongoing (n) press prevent the use of prior (n − 1) duration information? Light stimulation continued up until the lever press was terminated (Figure 4B) and stimulation occurred post task-acquisition (Figure 4C).

Both ChR2 and YFP mice reached similar rates of lever pressing and performance throughout training, including the 5 daily sessions during which light was delivered (Figures 4D and 4F, $ps > 0.05$), suggesting that acute disruptions of lOFC activity during action execution on a small subset of lever presses did not affect gross performance measures. A behavioral LME model that accounted for the presence or absence of the excitatory opsin in each animal found a significant interaction between the presence of the opsin and the predictive relationship between current and prior lever press durations ($Duration_{n-1}$ * Treatment) (Table S5, top, $p < 0.001$). To determine the direct effects of stimulation on the predictive relationship between prior (n − 1) and ongoing (n) lever press durations, we built post-hoc LME models using either the ChR2 or YFP datasets that accounted for the presence or absence of light stimulation in each lever press. We found a significant

interaction between prior press stimulation and the predictive relationship between current and prior press durations ($Duration_{n-1}$ * $Stimulation_{n-1}$) in ChR2-expressing mice that was absent in fluorophore control mice (Table S5, middle and bottom). Inspection of the LME model interaction β coefficients in ChR2 mice data showed light activation during the n − 1 lever press reduced the contribution of n−1 lever press duration from informing the ongoing (n) lever press ($p < 0.05$) (Figure 4G; left and right bar comparison). However, light activation during the ongoing (n) lever press did not reduce that lever press's reliance on n − 1 duration information ($p > 0.05$) (Figure 4G; middle and right bar comparison). These data suggest the proper patterning of lOFC activity supports processes related to the encoding (i.e. significant effect on n − 1 press activation), but not the retrieval and use (i.e. non-significant effect on n press activation) of action information to guide future action execution. Light activation did not induce selective decreases or increases to the lever press duration itself (Figure 4H; $p > 0.05$). In addition, light activation did not change the proportion of successful lever presses (Figure S3A) nor the time it took to initiate the subsequent lever press (Figure S3B). In conjunction with our PV+ inhibitory population disruptions, these data suggest proper patterning of lOFC activity supports processes related to the encoding of action information to guide future action execution.

## Loss of functional lOFC circuit increases reliance on immediate prior actions and outcomes

Different action strategies can be used to achieve the same goal.[79,1,80,3] When one circuit is offline another may be recruited to support decision-making and adaptive behavior (i.e the emitted behavior does not reflect the function of the perturbed circuit).[57,40,58,59] Within this framework, chronic removal of OFC circuits, as is often done in lesion studies, could bias recruitment of compensatory or parallel mechanisms for volitional action control. To test how chronic lesions to lOFC projection neurons would impact lever-press hold down task performance, prior to training we bilaterally ablated OFC CamKII+ neurons using a cre-dependent caspase approach that committed infected neurons to apoptosis (rAAV5/AAV-Flex-taCasP3-TEVP) (Figures 5A, 5B and S4A).[81,82]

LOFC lesions improved task efficiency. lOFC lesioned mice had lower overall response rates than Sham mice during 1600 ms training (Figure 5C; two-way RM ANOVA, main effect of Session $F_{1.912,}$ 70.73 = 16.34, $p < 0.0001$ and an interaction (Session * Treatment) $F_{4, 148} = 2.875$, $p < 0.05$). However, lesioned mice showed a higher percentage of rewarded lever presses than Sham mice (Figure 5D; two-way RM ANOVA, main effect of Session $F_{2.347, 86.83} = 22.55$, $p < 0.0001$, and Treatment $F_{1, 37} = 6.804$, $p < 0.05$). Furthermore, lOFC-lesioned mice showed a rightward shift in the distribution of lever press durations (Figure 5E; two-way RM ANOVA, main effect of Duration Bin $F_{2.872, 106.2} = 98.52$, $p < 0.0001$ and an interaction (Duration Bin * Treatment) $F_{9, 333} = 2.336$, $p < 0.05$) throughout the 1600 ms duration criterion sessions. A behavioral LME model that accounted for the presence or absence of the lesion in each animal found a significant interaction with treatment group reflecting an altered relationship between the current (n) and prior (n − 1) lever press durations ($Duration_{n-1}$ * Treatment) (Table S6; $p < 0.001$). Treatment group-segmented β coefficients showed a larger positive relationship between prior and subsequent durations in Lesion animals compared to Sham controls (significant compared to 1000

group-shuffled data) (Figure 5F, $p < 0.001$). We also found a significant interaction between the outcome of the prior lever press and lesion group ($Outcome_{n-1}$ * Treatment) (Table S6; $p = 0.001$). A representation of treatment group-segmented β coefficients showed a greater negative relationship between prior outcome and subsequent durations in Lesion mice compared to Sham mice (significant when compared to 1000 group-shuffled data) (Figure 5G; $p < 0.001$). The effect of lesions on use of prior action and outcome information was strongest during early 800 ms and early 1600 ms training (Figures S4B–S4E), suggesting that OFC lesions led to recruitment of other circuits which relied on immediate action and outcome history to a greater extent when contingencies were increased.

lOFC lesions have been reported to reduce behavioral flexibility and impair sensitivity to rule reversals.[83,84,17,23,85,46] We conducted an additional 5 daily sessions in which the duration criterion was reduced to 400 ms for a subset of animals (C57BL/6J; Lesion n = 16, 13 males, 3 females; Sham n = 8, 6 males, 2 females). We found that while Sham and Lesion mice showed similar rates of lever pressing (Figure 5H; two-way RM ANOVA, main effect of Session only $F_{1.645, 36.20} = 4.274$, $p < 0.05$), Lesion mice performed more efficiently than Sham mice as indexed by a higher percentage of rewarded lever presses (Figure 5I; two-way RM ANOVA, main effect of Treatment only $F_{1, 22} = 1.590$, $p < 0.01$) and showed a rightward shift in the distribution of lever press durations (Figure 5J; two-way RM ANOVA, main effect of Duration Bin $F_{2.369, 52.13} = 70.97$, $p < 0.0001$ and an interaction (Duration Bin * Treatment) $F_{9, 198} = 3.164$, $p < 0.01$). The above suggests mice with lOFC lesions did not adjust their performance to the same degree as sham animals when the duration contingency was reduced in duration. Instead, lOFC lesion mice continued to perform longer lever presses, a strategy that improved efficiency but differed from the exploration of effort that intact mice exhibited.

## Discussion

Here we identify a novel role for lOFC in action control. By examining continuous adjustments to volitional actions, we were able to separate control processes dictated by inferences about prior actions from those dictated by inferences of action-outcome contingency or expected outcome value. In doing so, we saw clear evidence that mice recruit and use lOFC activity to encode action-related information that can be used for inferences critical to adaptive control of behavior. A loss of lOFC circuits left mice more reliant on a strategy of repeating action execution to gain reward and impaired action exploration and the updating of action contingencies. This raises the hypothesis that lOFC circuit disruptions seen in psychiatric disorders may give way to compensatory mechanisms that promote repetitive action control by exploiting the reliance on learned rules, even when disadvantageous.

There is increasing evidence of lOFC disruption in psychiatric disorders characterized by disrupted action control, including substance use disorders and compulsive disorders,[86–89] highlighting the need for a greater understanding of OFC's contribution to the use of action-related information. Actions made during decision-making are often autonomous and unconstrained, occurring in contexts in which contingencies and associative structure of ongoing tasks are partially observable at best.[3,4,36,90] While prior single unit recordings

from largely unclassified populations have shown lOFC neurons can reflect sensory, predictive, and outcome-related information,[48,21,20,22] here we find that lOFC populations appear to be differentially recruited to support encoding of action-related information. These lOFC activity patterns were reminiscent of prior single-unit recording activity observations in rodents performing the same task,[42] suggesting single neuron population activity likely tracks population calcium activity. Furthermore, lOFC excitatory projection neuron activity reflected current and prior action-related information during ongoing action execution. Temporally precise and behavioral dependent perturbation to these endogenous lOFC CamKII+ neuron activity patterns decreased reliance on action-related information. Intriguingly, these lOFC CamKII+ activity patterns differed in timing and magnitude compared to PV+ interneuron activity patterns. PV+ populations showed relatively little performance-dependent modulation prior to and throughout action execution, little outcome encoding during reward checking behaviors, and maintained little representation of action-related information. While the use of population calcium measurements may not capture individual neuron encoding of action-related information, these findings do suggest that action-related information is reflected in recruitment of lOFC CamKII+ populations and to a much lesser extent PV+ populations. Cortical PV+interneurons are thought to gate information flow within cortical microcircuits through spike-timing enforcement of projection neuron firing.[68–71] Perhaps the observed differential patterns of lOFC activity are reflective of local microcircuit interactions that facilitate the flow of action-related information through this region.

OFC has been hypothesized to integrate information about prior experiences from broader circuits and relay such information to downstream targets in support of decision-making processes.[6,10,16,33,45,91,92] However, the loss of lOFC CamKII+ projection neuron populations did not lead to a loss of efficacy in performance or action control. Instead, lesioned mice showed more efficacious performance and persisted in making longer lever presses despite the change to a shorter duration criterion for success (Figure 5). Our results suggest a nuanced view of what lOFC may contribute to action control. While lOFC may not be necessary for direct action control per se, it does appear to be recruited when behavioral control recruits the inclusion of action and outcome-related inferences, broad experiences, and exploration.[38] Perhaps a functional loss of OFC circuits engaged compensatory mechanisms (e.g., recruitment of other circuits) that biased control of behavior to rely on more immediate sources of reward-related information.[79,93,80,94,95] In other words, lesioned animals may not have favored a shift in lever pressing strategy since long durations were still producing rewards. In addition, lesioned mice may have reduced the degree of exploration normally exhibited. Both hypotheses suggest lOFC CamKII+ projection neuron lesions left mice repeating actions to exploit a known rule. As task-related information has been seen broadly across the cortex,[96] how specific experiences recruit functional cortical and subcortical circuit activity should be explored in the future.

OFC dysfunction is found in disease states associated with repetitive behaviors and disrupted action control, such as in obsessive compulsive disorder and substance use disorders.[88,89] Investigating how information derived from past behaviors are integrated in OFC to influence subsequent actions can aid our understanding of how substances of abuse are sought out and consumed based on prior experience. In humans, repeated transcranial

magnetic stimulation studies targeting downstream OFC activity have been shown to be effective at reducing compulsivity.[97,98] Here we establish that lOFC population-specific activity can encode action-related information to influence future action implementation. While OFC neurons have been shown to have less action-related recruitment and activity modulation during motor responding compared to some other cortical areas,[99] discounting its role in processing action information in its entirety limits much needed investigations. The prior experimental discord over whether OFC contributes to action control may have arisen from the use of task parameters that were unable to isolate processes underlying action control from their relationship with associated outcomes or from examining choice behaviors that can use readily observable information. Such tasks may have also elicited a training-induced bias in recruitment of alternative mechanisms for action control. Thus, our findings support the hypothesis that lOFC circuits contribute to adaptive behaviors that rely on prior experience, and that their disruption may lead to an alteration of volitional action control biased towards excessive repetition.

## STAR Methods

### Resource availability

**Lead contact—**Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Christina M. Gremel, Ph.D. (cgremel@ucsd.edu).

**Materials availability—**This study did not generate new unique reagents.

#### Data and Code Availability

- The data reported in this paper has been deposited at https://figshare.com/articles/dataset/Orbitofrontal_cortex_populations_are_differentially_recruited_to_support_actions_/20997805 as a .mat file and is publicly available online as of the date of publication online and listed in the key resources table.
- All original code has been deposited at GitHub and is publicly available as of the date of publication online in the link listed in the key resources table. All scripts/functions were executed using Matlab 2019a.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

### Experimental model and subject details

C57BL/6J (n = 83, 58 males, 25 females) and $PV^{cre}$ ($Pvalb^{tm1(cre)Arbr}$: n = 36, 26 males, 10 females) mice (>7 weeks/50 PND) (The Jackson Laboratory, Bar Harbour, ME) were housed two to five per cage under a 14:10 hour light:dark in a temperature- and humidity-controlled room and had access to water *ad libitum*. Prior to behavioral procedures, mice were food restricted to 85–90% of their baseline weight for at least 2 days, and were fed a minimum of one hour after daily training (Labdiet 5015). Exploratory analyses for sex and genotype differences in the behavioral cohort revealed similar levels of behavioral performance, and

thus data was collapsed across sex and across genotype. Mice were at least 6 weeks of age prior to surgical procedures. All experiments were approved by the University of California San Diego Institutional Animal Care and Use Committee and were carried out in accordance with the National Institutes of Health (NIH) “Principles of Laboratory Care”. Investigators were not blind to the experimental groups. The Animal Care and Use Committee of the University of California, San Diego approved all experiments and experiments were conducted according to the National Institutes of Health (NIH) “Principles of Laboratory Care” guidelines.

## Method Details

**Behavioral Procedures—**Daily mouse training sessions occurred within sound attenuating operant chambers (Med-Associates, St Albans, VT) where lever presses (location counterbalanced, either left or right of the food magazine) were required for a reward outcome of regular ‘chow’ pellets (20 mg pellet per reinforcer, Bio-Serv formula F0071). On the first day of pre-training, mice were trained to retrieve pellets from the food magazine (no levers present) on a random time (RT) schedule, with a pellet outcome delivered on average every 120 seconds for 60 minutes. For the next 3 days of pre-training, lever presses were rewarded on a continuous reinforcement (CRF) schedule for up to 15 (CRF15), 30 (CRF30) or 60 (CRF60) pellet reward deliveries or until 90 minutes had passed. For surgical implant experiments, an additional CRF60 training day (for a total of 4 CRF days) was administered with the implant connected to habituate the animal to the tethered connection. Before each session in which the animal was tethered to a fiber optic cable, mice were exposed to a brief (< 60 seconds) bout of low-dose isoflurane anesthesia to connect the ferrule implant. To avoid confounding effects of anesthesia on brain activity, mice were then moved into the procedure room and monitored for a minimum of 30 min before placing them in the operant chamber and initiating the session. The start of each session triggered house-light illumination and the extension of the lever unless stated otherwise.

Following pre-training, mice were introduced to the hold down task. Lever presses now had a duration requirement, such that mice had to continue holding down the lever press for a fixed minimum amount of time in order to earn a pellet reward. Reward delivery occurred only after the termination of a lever press that exceeded the session’s minimum duration criteria, which began with > 800 ms for 5 daily sessions, followed by > 1600 ms for another 5 daily sessions. Each session ended when 90 minutes had elapsed or the mouse had earned 60 total reinforcers, at which point the house light turned off and the lever was retracted. Each lever press onset and termination was timestamped at a 20 ms time resolution to calculate its duration, along with pellet delivery and the start and end of head entries into the food magazine.

**Outcome Devaluation Testing—**A subset of the behavioral cohort (n = 18, 10 males, 8 females) was habituated to a novel cage and 20% sucrose solution for 1 hour each day. Following the last day of hold down training, we performed sensory-specific satiation across 2 consecutive days, consisting of counterbalanced valued and devalued days. For the valued day, the mice were allowed to freely consume 20% sucrose solution for 1 hour.

For the devalued day, mice were allowed to freely consume for 1 hour the pellet outcome previously earned in the lever press hold down task. One mouse that did not consume enough pellets (< 0.1 g) or sucrose (< 0.1 ml) during this free-access period was excluded from subsequent analysis (giving final n = 17, 10 males). Immediately following the feeding period, mice were placed into their respective operant chamber for a 10 minute session during which the number and duration of lever presses made were recorded, but no pellet reward was delivered. Investigators were not blind to the experimental groups. Response rate comparisons between valued and devalued days were made by normalizing each mouse's test day response rate (RR) to the average response rate of their corresponding last 2 days of hold down training using the following formula:

$$RRTest\ Day \div mean(RR1600ms4 + RR1600ms5)$$

**Surgical Procedures—**Mice first underwent isoflurane anesthesia (1–2%) before stereotaxic-guided intracranial injections via 500 nl volume Hamilton syringes (Reno, NV). Viral vectors were infused at a rate of 100 nl/minute and the syringe was then left unperturbed for 5 minutes to allow for diffusion after delivery. Mice were allowed to recover for a minimum of two weeks before the start of behavioral procedures. At the end of behavioral procedures, mice were euthanized and their brains were extracted and fixed in 4% paraformaldehyde. Optic fiber placement and viral expression was qualified by examining tracts in 50- to 100-μm-thick brain slices under a macro fluorescence microscope (Olympus MVX10). All surgical and behavioral experiments were performed during the light portion of the cycle.

For fiber photometry experiments, OFC was unilaterally targeted for viral injections at the following stereotaxic coordinates from Bregma: AP+2.7mm, L+1.65mm and V −2.6mm, with optic fiber ferrule placed V −2.5mm from the skull. For C57BL/6J mouse experiments, n = 9 mice (n = 6 males, n = 3 females) were injected with 300 nl of rAAV5/PAAV-CaMKIIa-GCaMP6s to express GCaMP6s under control of the $Ca^{2+}$ calmodulin dependent protein kinase IIα (CamKIIα) promoter. For $PV^{cre}$ mouse experiments, n = 8 mice (n = 5 males, n = 3 females) were injected with 300 nl of rAAV5/pAAV.CAG.Flex.GCaMP6s.WPRE.SV40 to express GCaMP6s via a Cre-dependent CAG promoter in PV+ neurons. An additional bilateral craniotomy was made over the posterior cerebellum for placement of screws to anchor a dental cement enclosure at the base of the ferrule to the base of the skull.

For optogenetic experiments, OFC was bilaterally targeted for viral injections at the following stereotaxic coordinates from Bregma: AP+2.6mm, L+1.75mm and V −2.1mm, with optic fiber ferrules placed V −1.9mm from the skull at a +12 degree orientation. For C57BL/6J mouse experiments, n = 12 mice (n = 7 males, n = 5 females) were injected with 250 nl of rAAV5/CamKII-hChR2(H134R)-eYFP-WPRE to express ChR2 under the CaMKIIα promoter for optogenetic activation or a combination of 250 nL of rAAV5/Ef1a-DIO-EYFP and 250 nl of rAAV5/CamKII-GFP-Cre for CamKIIα promoter fluorophore controls. For $PV^{cer}$ mouse experiments, n = 14 mice (n = 10 males, n = 4 females) were injected with 250 nl of rAAV5/Ef1a-DIO-hChR2(H134R)-eYFP to express cre-dependent

ChR2 in PV+ neurons for optogenetic activation or 250 nL of rAAV5/Ef1a-DIO-EYFP for fluorophore controls.

For lesion experiments, OFC was bilaterally targeted for viral injections at the following stereotaxic coordinates from Bregma: AP+2.7mm, L+1.65mm and V −2.6mm. C57BL/6J mice (n = 39, n = 27 males, n = 12 females) were injected with a combination of 250 nl of rAAV5/Ef1a-DIO-mCherry and 250 nL of rAAV5/AAV-Flex-taCasP3-TEVP for cre-dependent apoptosis lesions or 250 nl of rAAV5/Ef1a-DIO-mCherry for sham lesion controls. To assess the presence and spread of lesions, brains were first cut into 50 um slices and store at 4C in .1% sodium azide PBS before undergoing NeuN staining procedures using Alexa Fluor 488 Conjugate ABN78A4 Anti-NeuN (rabbit) antibody (Sigma-Aldrich). Slices were washed 3 times for 10 minutes with 1x PBS and pre-incubated in 10% Horse Serum and 0.3% Triton-X-100-PBS with 1% BSA for 1 hour. After, slices were incubated for 48 hours at 4C with primary antibody (1:500) in 2% horse serum and 0.3% Triton X-100-PBS-1% BSA 2%. Slices were then washed for 10 minutes with 3x PBS and stored at 4C until imaging.

**Fiber Photometry—**After pre-training procedures, ferrule-implanted animals were unilaterally attached to bifurcated 400 um optical fiber tethers (Thorlabs, Newton, NJ) through which a 470nm LED (Thorlabs, Newton, NJ) excited virally expressed GCaMP6s (< 70 μW/mm2). Emitted fluorescence was monitored through the core of the bifurcated fiber using a 4x objective (Olympus, Shinjuku, Japan) focused onto a CMOS camera (FLIR Systems, Wilsonville, OR). Regions of interest demarcating each fiber fork were created within the fiber core using Bonsai software[100] through which fluorescence intensity was captured at 20 Hz to produce two digitized signals, one for each animal connected to the bifurcated fiber. Analog behavioral timestamps for the beginning and end of each lever press, head entry, and reinforcer delivery periods were simultaneously sent to Bonsai software via TTL Med-PC pulses using microprocessors (Arduino Duo, from Arduino, Sumerville, MA) containing custom code. After each session, Bonsai software saved photometry signals and behavioral timestamps within comma-separated value files (.csv) that were then imported into Matlab (Mathworks Inc., Natick, MA) for subsequent analysis using custom scripts (see Code Availability). Raw fluorescence intensity signals underwent running median (5th order) and low pass (high cutoff frequency of 1 Hz) filtering to reduce noise and electrical artifacts. To correct for photobleaching in which a signal captured from fluorophores degrades by continuous light exposure during the session, we high pass filtered the signal with a low cutoff frequency of 0.001Hz. Filtered fluorescence intensity signals subsequently underwent a quality check for low expression and fiber decoupling. Briefly, sessions that did not exceed a 15 second moving window calculation of the signal's 97.5 percentile by a minimum 1% fluorescence change[101] or did not pass a visual inspection for within-session fiber-ferrule decoupling artifacts were excluded from further analyses. Peri-event changes in fluorescence intensity were then calculated via z-score normalization to each corresponding pre-lever press onset period (i.e. −5 seconds to −2 seconds prior to lever press). These z-scored fluorescence traces were then combined across all mice within a group to preserve the variance seen within a subject. Activity during the ongoing lever press duration was modified using Akima interpolation via MATLAB's *interp1* function,

excluding any lever press that was fewer than 2 samples (i.e. 100 ms) in duration that would invalidate interpolation. Comparisons between rewarded and unrewarded lever press traces were made using running permutation tests (1000 shuffles) that required at least 5 consecutive samples (or 3 consecutive samples for interpolated activity) to be different from one another.[77] Population $Ca^{2+}$ activity traces were then smoothed with MATLAB's Savitzky–Golay *smoothdata* method using a 20 sample (or 1 sample for interpolated activity) sliding window for visual display purposes only.

**Optogenetic Excitation—**Optogenetic excitation occurred only in the additional 5 sessions (days 6 to 10) during which the minimum duration criteria was 1600 ms. LEDs (470nm, Thorlabs) used for optogenetic excitation experiments were triggered by TTL pulses emitted from Med-PC operant chambers via Arduino Duos programmed with custom code. Sheathed (200 uM) optic fiber cables were coupled to bilaterally implanted ferrules (>= 1mW output at ferrule tip) through which a closed-loop system delivered light at 20 Hz (5 ms pulses) throughout the entirety of each (or every 7th) lever press duration.

**Linear Mixed Effects Models of Behavior—**Linear Mixed Effects (LME) models were built to investigate the predictive relationship between the duration of individual lever presses (n) and the lever press occurring immediately prior to it (n − 1)[33]. Random intercept terms for mouse and training day were included to account for the repeated, non-independent structure of the aggregated session data. To account for variance explained by the overall performance within a session, fixed terms included the overall percentage of rewarded lever presses as well as the timestamps of each lever press. To test how predictive relationships were contingent upon their sequential order, beta coefficient outputs pertaining to each behavioral measurement of interest were compared to a 1000 order shuffled (unless otherwise specified) distribution of beta coefficients using permutation testing (Table S1, top). Importantly, shuffling occurred within individual sessions/mice to preserve overall performance statistics (e.g. total lever presses made), and the order shuffling for each behavioral covariate occurred independently from each other. Thus the LME model for the behavioral cohort consisted of the following formula:

$$D_n = \beta_0 + \beta_D D_{n-1} + \beta_O O_{n-1} + \beta_{HE} HE_{n-1} + \beta_{IPI} IPI_{n-1} + \beta_t(t) + \beta_{\%}(\%) + (1 \mid M) + (1 \mid D) + \varepsilon_i$$

Where $D_n$ is the current lever press duration, $D_{n-1}$ is the prior lever press duration (in ms), $O_{n-1}$ is the outcome of the prior lever press (binary 1 for reward, 0 for no reward), $HE_{n-1}$ is the indicator of whether a head entry was made between the current and prior lever press (binary 1 for head entry made, 0 for no head entry made), $IPI_{n-1}$ is the interpress interval (in ms), and $B_x$ is the linear regression coefficient for each corresponding behavioral covariate term × ($\beta_0$ being the intercept term). Covariates for lever press timestamps (t, in ms) and overall percentage of rewarded lever presses (%) were included alongside random intercept terms for mouse (M) and day (D).

To determine how far back the predictive relationship existed between press n and any particular n-back press, we built and 100 shuffled-order tested a similar LME model that

included additional variables accounting for the duration of lever press n and n-back (n − 1 through n − 10) lever press durations as follows (Table S1, middle):

$$n = \beta_0 + \beta_{n-1}n_{-1} + \beta_{n-2}n_{-2} + \ldots + \beta_{n-10}n_{-10} + \beta_t(t) + \beta_{\%}(\%) + (1 \mid M) + (1 \mid D) + \varepsilon_i$$

We built and 100 shuffled-order tested an additional LME model that included interaction terms to determine how prior behavioral variables (i.e. prior reward, checking, and interpress interval) compounded their effect on the subsequent lever press duration (Table S1, bottom):

$$D_n = \beta_0 + \beta_D D_{n-1} + \beta_O O_{n-1} + \beta_{HE} HE_{n-1} + \beta_{IPI} IPI_{n-1} + \beta_{O*D} O_{n-1} * D_{n-1} \beta_{HE*D} H_{n-1} * D_{n-1} + \beta_{IPI*D} IPI_{n-1} * D_{n-1} + \beta_t(t) + \beta_{\%}(\%) + (1 \mid M) + (1 \mid D) + \varepsilon_i$$

Regression coefficient terms $\beta_x$ and shuffled-order testing procedures were as previously described, with the added covariates for main effects of prior duration ($D_{n-1}$, lever press duration in ms) and its interactions with prior lever press outcome ($O_{n-1}$*$D_{n-1}$), prior presence of a head entry ($HE_{n-1}$*$D_{n-1}$), and interpress interval ($IPI_{n-1}$*$D_{n-1}$).

To determine how the predictive relationship of behavioral covariates for current lever press durations were affected by experimental manipulations (e.g. optogenetic excitation via ChR2), we built and 1000 shuffled-order tested similar LME models that included additional variables accounting for treatment group main effects and interactions as follows (Tables S4–S6):

$$D_n = \beta_0 + \beta_D D_{n-1} + \beta_O O_{n-1} + \beta_{HE} HE_{n-1} + \beta_{IPI} IPI_{n-1} + \beta_t(t) + \beta_{\%}(\%) + \beta_{Tx}(Tx) + \beta_{D*Tx} D_{n\text{-}1} * Tx + \beta_{O*Tx} O_{n-1} * Tx + \beta_{HE*Tx} HE_{n-1} * Tx + \beta_{IPI*Tx} IPI_{n-1} * Tx + (1|M) + (1|D) + \varepsilon_i$$

Regression coefficient terms $\beta_x$ in these models were as previously described, with the added covariates for main effects of treatment (Tx, binary 1 for experimental and 0 for control groups) and its interactions with prior lever press duration ($D_{n-1}$*Tx), prior outcome ($O_{n-1}$*Tx), prior presence of a head entry ($HE_{n-1}$*Tx), and interpress interval ($IPI_{n-1}$*Tx). For the optogenetic excitation experiment in which only every 7th lever press triggered light delivery, a post-hoc LME model was tested using only the ChR2 group. In this case, however, the treatment main effect term Tx (and associated interaction terms) instead indicated the presence or absence of optogenetic stimulation for each individual lever press.

**Linear Mixed Effects Models of $Ca^{2+}$ Activity—**For OFC $Ca^{2+}$ fluorescence activity monitoring experiments, LME models were built to predict $Ca^{2+}$ activity within peri-event epochs given current and prior lever press durations alongside other behavioral variables. For these models, only data collected from the 1600 ms minimum duration criterion days were used. The mean area under the curve of activity traces during each of three epochs (−1s to 0s before lever press onset, lever press duration, and 0s to +1s after lever press release) was calculated to predict activity at each of these three time points using the following formula (Tables S2 and S3):

$$A_n = \beta_0 + \beta_D D_n + \beta_{D\text{-}1} D_{n-1} + \beta_O O_{n-1} + \beta_{HE} HE_{n-1} + \beta_{IPI} IPI_{n-1} + \beta_t (t) + \beta_{\%} (\%) + \beta_A A_{n-1} + (1|M) + (1|D) + \varepsilon_i$$

Where $A_n$ is $Ca^{2+}$ activity associated with the current lever press epoch (pre-onset, duration, or post-offset). Regression coefficient terms $\beta_x$ in these models were as previously described, with the added covariates for main effects of current duration ($D_n$, in ms) and prior lever press activity ($A_{n-1}$) during that epoch to control for $Ca^{2+}$ activity autocorrelation. LME regression coefficients for behavior measures of interest were compared to 1000 order shuffled datasets to test whether their predictive ability was due to the subsequent relationship.

### Quantification and Statistical Analysis

All analyses were two-tailed and statistical significance was defined as an $\alpha$ of $p < 0.05$. Statistical analysis was performed using GraphPad Prism 8.3.0 (GraphPad Software) and custom MATLAB R2019a (MathWorks) scripts using a PC desktop with Windows 10. Acquisition data, including lever presses, response rate, and percentage of lever presses that were rewarded were analyzed using one-way or two-way repeated measures ANOVAs with Greenhouse-Geisser corrections and Šidák corrections for post-hoc multiple comparisons unless otherwise noted. For outcome devaluation testing, a paired parametric t-test was performed to examine whether sensory-specific satiety reduced lever press responses on the devalued day compared to the valued day. For each LME model, we report the average regression coefficient ($\beta$), which measures the effect size and indicates how much a change in a predictor variable will change the output (e.g. lever press duration). Unless stated otherwise, significant predictors underwent follow-up permutation test comparisons for $\beta$ coefficient values against a distribution of 1000 order or group shuffled versions of the same variable. For $Ca^{2+}$ activity comparisons (i.e. reward vs no reward), permutation testing required 5 consecutive samples (or 3 consecutive samples for interpolated activity) that passed the threshold for significance. Lever presses longer than 10 seconds were excluded from all $Ca^{2+}$ activity analyses. Data are presented as mean ± SEM.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

This work was funded by F99-NS120434 (C.C.), F31AA027439 (D.C.S.), R01AA026077 (C.M.G.), and a Whitehall Foundation Award (C.M.G). Some figures included schematics created with BioRender.com.

## References

1. Balleine BW, and Dickinson A (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37, 407–419. 10.1016/s0028-3908(98)00033-1. [PubMed: 9704982]
2. Bouton ME, and Balleine BW (2019). Prediction and control of operant behavior: What you see is not all there is. Behav. Anal. Wash. DC 19, 202–212. 10.1037/bar0000108.

3. Balleine BW (2019). The Meaning of Behavior: Discriminating Reflex and Volition in the Brain. Neuron 104, 47–62. 10.1016/j.neuron.2019.09.024. [PubMed: 31600515]
4. Yoo SBM, Hayden BY, and Pearson JM (2021). Continuous decisions. Philos. Trans. R. Soc. B Biol. Sci. 376, 20190664. 10.1098/rstb.2019.0664.
5. Wilson RC, Takahashi YK, Schoenbaum G, and Niv Y (2014). Orbitofrontal cortex as a cognitive map of task space. Neuron 81, 267–279. 10.1016/j.neuron.2013.11.005. [PubMed: 24462094]
6. Schuck NW, Cai MB, Wilson RC, and Niv Y (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. Neuron 91, 1402–1412. 10.1016/j.neuron.2016.08.019. [PubMed: 27657452]
7. Wikenheiser AM, and Schoenbaum G (2016). Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. Nat. Rev. Neurosci. 17, 513–523. 10.1038/nrn.2016.56. [PubMed: 27256552]
8. Lopatina N, Sadacca BF, McDannald MA, Styer CV, Peterson JF, Cheer JF, and Schoenbaum G (2017). Ensembles in medial and lateral orbitofrontal cortex construct cognitive maps emphasizing different features of the behavioral landscape. Behav. Neurosci. 131, 201–212. 10.1037/bne0000195. [PubMed: 28541078]
9. Sadacca BF, Wied HM, Lopatina N, Saini GK, Nemirovsky D, and Schoenbaum G (2018). Orbitofrontal neurons signal sensory associations underlying model-based inference in a sensory preconditioning task. eLife 7, e30373. 10.7554/eLife.30373. [PubMed: 29513220]
10. Niv Y (2019). Learning task-state representations. Nat. Neurosci. 22, 1544–1553. 10.1038/s41593-019-0470-8. [PubMed: 31551597]
11. Gardner MPH, and Schoenbaum G (2021). The orbitofrontal cartographer. Behav. Neurosci. 135, 267–276. 10.1037/bne0000463. [PubMed: 34060879]
12. Gallagher M, McMahan RW, and Schoenbaum G (1999). Orbitofrontal Cortex and Representation of Incentive Value in Associative Learning. J. Neurosci. 19, 6610–6614. 10.1523/JNEUROSCI.19-15-06610.1999. [PubMed: 10414988]
13. Ostlund SB, and Balleine BW (2007). Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J. Neurosci. Off. J. Soc. Neurosci. 27, 4819–4825. 10.1523/JNEUROSCI.5443-06.2007.
14. Burke KA, Takahashi YK, Correll J, Brown PL, and Schoenbaum G (2009). Orbitofrontal inactivation impairs reversal of Pavlovian learning by interfering with disinhibition of responding for previously unrewarded cues. Eur. J. Neurosci. 30, 1941–1946. 10.1111/j.1460-9568.2009.06992.x. [PubMed: 19912335]
15. Morrison SE, and Salzman CD (2011). Representations of appetitive and aversive information in the primate orbitofrontal cortex. Ann. N. Y. Acad. Sci. 1239, 59–70. 10.1111/j.1749-6632.2011.06255.x. [PubMed: 22145876]
16. Namboodiri VMK, Otis JM, van Heeswijk K, Voets ES, Alghorazi RA, Rodriguez-Romaguera J, Mihalas S, and 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. 10.1038/s41593-019-0408-1. [PubMed: 31160741]
17. Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, and Gallagher M (2003). Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn. Mem. Cold Spring Harb. N 10, 129–140. 10.1101/lm.55203.
18. Roesch MR, Taylor AR, and Schoenbaum G (2006). Encoding of Time-Discounted Rewards in Orbitofrontal Cortex Is Independent of Value Representation. Neuron 51, 509–520. 10.1016/j.neuron.2006.06.027. [PubMed: 16908415]
19. Rudebeck PH, and Murray EA (2008). Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. J. Neurosci. Off. J. Soc. Neurosci. 28, 8338–8343. 10.1523/JNEUROSCI.2272-08.2008.
20. Nogueira R, Abolafia JM, Drugowitsch J, Balaguer-Ballester E, Sanchez-Vives MV, and Moreno-Bote R (2017). Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. Nat. Commun. 8, 1–13. 10.1038/ncomms14823. [PubMed: 28232747]

21. Riceberg JS, and Shapiro ML (2017). Orbitofrontal Cortex Signals Expected Outcomes with Predictive Codes When Stable Contingencies Promote the Integration of Reward History. J. Neurosci. 37, 2010–2021. 10.1523/JNEUROSCI.2951-16.2016. [PubMed: 28115481]
22. Hocker DL, Brody CD, Savin C, and Constantinople CM (2021). Subpopulations of neurons in lOFC encode previous and current rewards at time of choice. eLife 10, e70129. 10.7554/eLife.70129. [PubMed: 34693908]
23. Izquierdo A, Suda RK, and Murray EA (2004). Bilateral Orbital Prefrontal Cortex Lesions in Rhesus Monkeys Disrupt Choices Guided by Both Reward Value and Reward Contingency. J. Neurosci. 24, 7540–7548. 10.1523/JNEUROSCI.1921-04.2004. [PubMed: 15329401]
24. Burke KA, Franz TM, Miller DN, and Schoenbaum G (2008). The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. Nature 454, 340–344. 10.1038/nature06993. [PubMed: 18563088]
25. Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenzi A, and Schoenbaum G (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. Science 338, 953–956. 10.1126/science.1227489. [PubMed: 23162000]
26. Stalnaker TA, Cooch NK, McDannald MA, Liu T-L, Wied H, and Schoenbaum G (2014). Orbitofrontal neurons infer the value and identity of predicted outcomes. Nat. Commun. 5, 3926. 10.1038/ncomms4926. [PubMed: 24894805]
27. Bradfield LA, Dezfouli A, van Holstein M, Chieng B, and Balleine BW (2015). Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. Neuron 88, 1268–1280. 10.1016/j.neuron.2015.10.044. [PubMed: 26627312]
28. Rich EL, and Wallis JD (2016). Decoding subjective decisions from orbitofrontal cortex. Nat. Neurosci. 19, 973–980. 10.1038/nn.4320. [PubMed: 27273768]
29. Baltz ET, Yalcinbas EA, Renteria R, and Gremel CM (2018). Orbital frontal cortex updates state-induced value change for decision-making. eLife 7, e35988. 10.7554/eLife.35988. [PubMed: 29897332]
30. Malvaez M, Shieh C, Murphy MD, Greenfield VY, and Wassum KM (2019). Distinct cortical-amygdala projections drive reward value encoding and retrieval. Nat. Neurosci. 22, 762–769. 10.1038/s41593-019-0374-7. [PubMed: 30962632]
31. Klaus A, Alves da Silva J, and Costa RM (2019). What, If, and When to Move: Basal Ganglia Circuits and Self-Paced Action Initiation. Annu. Rev. Neurosci. 42, 459–483. 10.1146/annurev-neuro-072116-031033. [PubMed: 31018098]
32. Schreiner DC, Yalcinbas EA, and Gremel CM (2021). A Push For Examining Subjective Experience in Value-Based Decision-Making. Curr. Opin. Behav. Sci. 41, 45–49. 10.1016/j.cobeha.2021.03.020. [PubMed: 34056054]
33. Schreiner DC, Cazares C, Renteria R, and Gremel CM (2022). Information normally considered task-irrelevant drives decision-making and affects premotor circuit recruitment. Nat. Commun. 13, 2134. 10.1038/s41467-022-29807-2. [PubMed: 35440120]
34. Yalcinbas EA, Cazares C, and Gremel CM (2021). Call for a more balanced approach to understanding orbital frontal cortex function. Behav. Neurosci. 135, 255–266. 10.1037/bne0000450. [PubMed: 34060878]
35. Haggard P (2008). Human volition: towards a neuroscience of will. Nat. Rev. Neurosci. 9, 934–946. 10.1038/nrn2497. [PubMed: 19020512]
36. Costa RM (2011). A selectionist account of de novo action learning. Curr. Opin. Neurobiol. 21, 579–586. 10.1016/j.conb.2011.05.004. [PubMed: 21641793]
37. Daw ND, O'Doherty JP, Dayan P, Seymour B, and Dolan RJ (2006). Cortical substrates for exploratory decisions in humans. Nature 441, 876–879. 10.1038/nature04766. [PubMed: 16778890]
38. Hogeveen J, Mullins TS, Romero JD, Eversole E, Rogge-Obando K, Mayer AR, and Costa VD (2022). The neurocomputational bases of explore-exploit decision-making. Neuron. 10.1016/j.neuron.2022.03.014.
39. Furuyashiki T, Holland PC, and Gallagher M (2008). Rat Orbitofrontal Cortex Separately Encodes Response and Outcome Information during Performance of Goal-Directed Behavior. J. Neurosci. 28, 5127–5138. 10.1523/JNEUROSCI.0319-08.2008. [PubMed: 18463266]

40. Gremel CM, and Costa RM (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun. 4, 2264. 10.1038/ncomms3264. [PubMed: 23921250]
41. Simon NW, Wood J, and Moghaddam B (2015). Action-outcome relationships are represented differently by medial prefrontal and orbitofrontal cortex neurons during action execution. J. Neurophysiol. 114, 3374–3385. 10.1152/jn.00884.2015. [PubMed: 26467523]
42. Cazares C, Schreiner DC, and Gremel CM (2021). Different Effects of Alcohol Exposure on Action and Outcome-Related Orbitofrontal Cortex Activity. eNeuro 8. 10.1523/ENEURO.0052-21.2021.
43. Gourley SL, Olevska A, Zimmermann KS, Ressler KJ, DiLeone RJ, and Taylor JR (2013). The orbitofrontal cortex regulates outcome-based decision-making via the lateral striatum. Eur. J. Neurosci. 38. 10.1111/ejn.12239.
44. Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, Deisseroth K, Lovinger DM, and Costa RM (2016). Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. Neuron 90, 1312–1324. 10.1016/j.neuron.2016.04.043. [PubMed: 27238866]
45. Renteria R, Baltz ET, and Gremel CM (2018). Chronic alcohol exposure disrupts top-down control over basal ganglia action selection to produce habits. Nat. Commun. 9, 211. 10.1038/s41467-017-02615-9. [PubMed: 29335427]
46. Rhodes SEV, and Murray EA (2013). Differential Effects of Amygdala, Orbital Prefrontal Cortex, and Prelimbic Cortex Lesions on Goal-Directed Behavior in Rhesus Macaques. J. Neurosci. 33, 3380–3389. 10.1523/JNEUROSCI.4374-12.2013. [PubMed: 23426666]
47. Parkes SL, Ravassard PM, Cerpa J-C, Wolff M, Ferreira G, and Coutureau E (2018). Insular and Ventrolateral Orbitofrontal Cortices Differentially Contribute to Goal-Directed Behavior in Rodents. Cereb. Cortex N. Y. N 1991 28, 2313–2325. 10.1093/cercor/bhx132.
48. Padoa-Schioppa C, and Assad JA (2006). Neurons in the orbitofrontal cortex encode economic value. Nature 441, 223–226. 10.1038/nature04676. [PubMed: 16633341]
49. Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, and Rushworth MFS (2008). Frontal Cortex Subregions Play Distinct Roles in Choices between Actions and Stimuli. J. Neurosci. 28, 13775–13785. 10.1523/JNEUROSCI.3541-08.2008. [PubMed: 19091968]
50. Camille N, Tsuchida A, and Fellows LK (2011). Double Dissociation of Stimulus-Value and Action-Value Learning in Humans with Orbitofrontal or Anterior Cingulate Cortex Damage. J. Neurosci. 31, 15048–15052. 10.1523/JNEUROSCI.3164-11.2011. [PubMed: 22016538]
51. Fellows LK (2011). Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. Ann. N. Y. Acad. Sci. 1239, 51–58. 10.1111/j.1749-6632.2011.06229.x. [PubMed: 22145875]
52. Luk C-H, and Wallis JD (2013). Choice Coding in Frontal Cortex during Stimulus-Guided or Action-Guided Decision-Making. J. Neurosci. 33, 1864–1871. 10.1523/JNEUROSCI.4920-12.2013. [PubMed: 23365226]
53. Cai X, and Padoa-Schioppa C (2014). Contributions of Orbitofrontal and Lateral Prefrontal Cortices to Economic Choice and the Good-to-action Transformation. Neuron 81, 1140–1151. 10.1016/j.neuron.2014.01.008. [PubMed: 24529981]
54. Panayi MC, and Killcross S (2018). Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. eLife 7, e37357. 10.7554/eLife.37357. [PubMed: 30044220]
55. Grattan LE, and Glimcher PW (2014). Absence of Spatial Tuning in the Orbitofrontal Cortex. PLOS ONE 9, e112750. 10.1371/journal.pone.0112750. [PubMed: 25386837]
56. Duan LY, Horst NK, Cranmore SAW, Horiguchi N, Cardinal RN, Roberts AC, and Robbins TW (2021). Controlling one's world: Identification of sub-regions of primate PFC underlying goal-directed behavior. Neuron 109, 2485–2498.e5. 10.1016/j.neuron.2021.06.003. [PubMed: 34171290]
57. Yin HH, Knowlton BJ, and Balleine BW (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav. Brain Res. 166, 189–196. 10.1016/j.bbr.2005.07.012. [PubMed: 16153716]

58. Li N, Daie K, Svoboda K, and Druckmann S (2016). Robust neuronal dynamics in premotor cortex during motor planning. Nature 532, 459–464. 10.1038/nature17643. [PubMed: 27074502]
59. Fresno V, Parkes SL, Faugère A, Coutureau E, and Wolff M (2019). A thalamocortical circuit for updating action-outcome associations. eLife 8, e46187. 10.7554/eLife.46187. [PubMed: 31012845]
60. Inagaki HK, Chen S, Ridder MC, Sah P, Li N, Yang Z, Hasanbegovic H, Gao Z, Gerfen CR, and Svoboda K (2022). A midbrain-thalamus-cortex circuit reorganizes cortical dynamics to initiate movement. Cell 185, 1065–1081.e23. 10.1016/j.cell.2022.02.006. [PubMed: 35245431]
61. Cavada C, Compañy T, Tejedor J, Cruz-Rizzolo RJ, and Reinoso-Suárez F (2000). The Anatomical Connections of the Macaque Monkey Orbitofrontal Cortex. A Review. Cereb. Cortex 10, 220–242. 10.1093/cercor/10.3.220. [PubMed: 10731218]
62. Rolls ET (2004). The functions of the orbitofrontal cortex. Brain Cogn. 55, 11–29. 10.1016/S0278-2626(03)00277-X. [PubMed: 15134840]
63. Price JL (2007). Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. Ann. N. Y. Acad. Sci. 1121, 54–71. 10.1196/annals.1401.008. [PubMed: 17698999]
64. Zhang S, Xu M, Chang W-C, Ma C, Hoang Do JP, Jeong D, Lei T, Fan JL, and Dan Y (2016). Organization of long-range inputs and outputs of frontal cortex for top-down control. Nat. Neurosci. 19, 1733–1742. 10.1038/nn.4417. [PubMed: 27749828]
65. Murphy MJM, and Deutch AY (2018). Organization of afferents to the orbitofrontal cortex in the rat. J. Comp. Neurol. 526, 1498–1526. 10.1002/cne.24424. [PubMed: 29524205]
66. Barreiros IV, Panayi MC, and Walton ME (2021). Organization of Afferents along the Anterior-posterior and Medial-lateral Axes of the Rat Orbitofrontal Cortex. Neuroscience 460, 53–68. 10.1016/j.neuroscience.2021.02.017. [PubMed: 33609638]
67. Sohal VS, Zhang F, Yizhar O, and Deisseroth K (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. Nature 459, 698–702. 10.1038/nature07991. [PubMed: 19396159]
68. Isaacson JS, and Scanziani M (2011). How Inhibition Shapes Cortical Activity. Neuron 72, 231–243. 10.1016/j.neuron.2011.09.027. [PubMed: 22017986]
69. Hu H, Gan J, and Jonas P (2014). Fast-spiking, parvalbumin+ GABAergic interneurons: From cellular design to microcircuit function. Science 345. 10.1126/science.1255263.
70. Kepecs A, and Fishell G (2014). Interneuron cell types are fit to function. Nature 505, 318–326. 10.1038/nature12983. [PubMed: 24429630]
71. Ferguson KA, and Cardin JA (2020). Mechanisms underlying gain modulation in the cortex. Nat. Rev. Neurosci. 21, 80–92. 10.1038/s41583-019-0253-y. [PubMed: 31911627]
72. Skinner BF (1938). The behavior of organisms: an experimental analysis (Appleton-Century).
73. Platt JR, Kuch DO, and Bitgood SC (1973). Rats' lever-press durations as psychophysical judgements of time. J. Exp. Anal. Behav. 19, 239–250. 10.1901/jeab.1973.19-239. [PubMed: 4716170]
74. Yin H (2009). The role of the murine motor cortex in action duration and order. Front. Integr. Neurosci. 3.
75. Fan D, Rossi MA, and Yin HH (2012). Mechanisms of action selection and timing in substantia nigra neurons. J. Neurosci. Off. J. Soc. Neurosci. 32, 5534–5548. 10.1523/JNEUROSCI.5924-11.2012.
76. Jovanovic V, Fishbein AR, de la Mothe L, Lee K-F, and Miller CT (2022). Behavioral context affects social signal representations within single primate prefrontal cortex neurons. Neuron 110, 1318–1326.e4. 10.1016/j.neuron.2022.01.020. [PubMed: 35108498]
77. Jean-Richard-dit-Bressel P, Clifford CWG, and McNally GP (2020). Analyzing Event-Related Transients: Confidence Intervals, Permutation Tests, and Consecutive Thresholds. Front. Mol. Neurosci. 13.
78. Li N, Chen S, Guo ZV, Chen H, Huo Y, Inagaki HK, Chen G, Davis C, Hansel D, Guo C, et al. (2019). Spatiotemporal constraints on optogenetic inactivation in cortical circuits. eLife 8, e48622. 10.7554/eLife.48622. [PubMed: 31736463]

79. Dickinson A (1985). Actions and Habits: The Development of Behavioural Autonomy. Philos. Trans. R. Soc. Lond. Ser. B 308, 67. 10.1098/rstb.1985.0010.
80. Balleine BW, and O'Doherty JP (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol. 35, 48–69. 10.1038/npp.2009.131.
81. Fink SL, and Cookson BT (2005). Apoptosis, Pyroptosis, and Necrosis: Mechanistic Description of Dead and Dying Eukaryotic Cells. Infect. Immun. 73, 1907–1916. 10.1128/IAI.73.4.1907-1916.2005. [PubMed: 15784530]
82. Yang CF, Chiang MC, Gray DC, Prabhakaran M, Alvarado M, Juntti SA, Unger EK, Wells JA, and Shah NM (2013). Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. Cell 153, 896–909. 10.1016/j.cell.2013.04.017. [PubMed: 23663785]
83. Bechara A, Damasio H, Damasio AR, and Lee GP (1999). Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making. J. Neurosci. 19, 5473–5481. 10.1523/JNEUROSCI.19-13-05473.1999. [PubMed: 10377356]
84. Schoenbaum G, Nugent SL, Saddoris MP, and Setlow B (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport 13, 885–890. 10.1097/00001756-200205070-00030. [PubMed: 11997707]
85. Stalnaker TA, Franz TM, Singh T, and Schoenbaum G (2007). Basolateral Amygdala Lesions Abolish Orbitofrontal-Dependent Reversal Impairments. Neuron 54, 51–58. 10.1016/j.neuron.2007.02.014. [PubMed: 17408577]
86. Milad MR, and Rauch SL (2012). Obsessive-compulsive disorder: beyond segregated cortico-striatal pathways. Trends Cogn. Sci. 16, 43–51. 10.1016/j.tics.2011.11.003. [PubMed: 22138231]
87. Pauls DL, Abramovitch A, Rauch SL, and Geller DA (2014). Obsessive-compulsive disorder: an integrative genetic and neurobiological perspective. Nat. Rev. Neurosci. 15, 410–424. 10.1038/nrn3746. [PubMed: 24840803]
88. Robbins TW, Vaghi MM, and Banca P (2019). Obsessive-Compulsive Disorder: Puzzles and Prospects. Neuron 102, 27–47. 10.1016/j.neuron.2019.01.046. [PubMed: 30946823]
89. Lüscher C, Robbins TW, and Everitt BJ (2020). The transition to compulsion in addiction. Nat. Rev. Neurosci. 21, 247–263. 10.1038/s41583-020-0289-z. [PubMed: 32231315]
90. Murakami M, Vicente MI, Costa GM, and Mainen ZF (2014). Neural antecedents of self-initiated actions in secondary motor cortex. Nat. Neurosci. 17, 1574–1582. 10.1038/nn.3826. [PubMed: 25262496]
91. Sias AC, Morse AK, Wang S, Greenfield VY, Goodpaster CM, Wrenn TM, Wikenheiser AM, Holley SM, Cepeda C, Levine MS, et al. (2021). A bidirectional corticoamygdala circuit for the encoding and retrieval of detailed reward memories. eLife 10, e68617. 10.7554/eLife.68617. [PubMed: 34142660]
92. Schreiner DC, and Gremel CM (2018). Orbital Frontal Cortex Projections to Secondary Motor Cortex Mediate Exploitation of Learned Rules. Sci. Rep. 8, 10979. 10.1038/s41598-018-29285-x. [PubMed: 30030509]
93. Daw ND, Niv Y, and Dayan P (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 8, 1704–1711. 10.1038/nn1560. [PubMed: 16286932]
94. Doll BB, Simon DA, and Daw ND (2012). The ubiquity of model-based reinforcement learning. Curr. Opin. Neurobiol. 22, 1075–1081. 10.1016/j.conb.2012.08.003. [PubMed: 22959354]
95. Drummond N, and Niv Y (2020). Model-based decision making and model-free learning. Curr. Biol. 30, R860–R865. 10.1016/j.cub.2020.06.051. [PubMed: 32750340]
96. Ren C, and Komiyama T (2021). Characterizing Cortex-Wide Dynamics with Wide-Field Calcium Imaging. J. Neurosci. 41, 4160–4168. 10.1523/JNEUROSCI.3003-20.2021. [PubMed: 33893217]
97. Nauczyciel C, Le Jeune F, Naudet F, Douabin S, Esquevin A, Vérin M, Dondaine T, Robert G, Drapier D, and Millet B (2014). Repetitive transcranial magnetic stimulation over the orbitofrontal cortex for obsessive-compulsive disorder: a double-blind, crossover study. Transl. Psychiatry 4, e436. 10.1038/tp.2014.62. [PubMed: 25203167]

98. Price RB, Gillan CM, Hanlon C, Ferrarelli F, Kim T, Karim HT, Renard M, Kaskie R, Degutis M, Wears A, et al. (2021). Effect of Experimental Manipulation of the Orbitofrontal Cortex on Short-Term Markers of Compulsive Behavior: A Theta Burst Stimulation Study. Am. J. Psychiatry 178, 459–468. 10.1176/appi.ajp.2020.20060821. [PubMed: 33726523]

99. Knudsen EB, and Wallis JD (2022). Taking stock of value in the orbitofrontal cortex. Nat. Rev. Neurosci. 23, 428–438. 10.1038/s41583-022-00589-2. [PubMed: 35468999]

100. Lopes G, Bonacchi N, Frazão J, Neto JP, Atallah BV, Soares S, Moreira L, Matias S, Itskov PM, Correia PA, et al. (2015). Bonsai: an event-based framework for processing and controlling data streams. Front. Neuroinformatics 9.

101. Markowitz JE, Gillis WF, Beron CC, Neufeld SQ, Robertson K, Bhagat ND, Peterson RE, Peterson E, Hyun M, Linderman SW, et al. (2018). The Striatum Organizes 3D Behavior via Moment-to-Moment Action Selection. Cell 174, 44–58.e17. 10.1016/j.cell.2018.04.019. [PubMed: 29779950]

## Highlights

- Lateral Orbitofrontal cortex (lOFC) populations represent action information.
- lOFC activity is important for encoding past action information for future use.
- lOFC activity is not important for use of prior action information.
- Chronic loss of lOFC activity recruits compensatory repetitive action control.

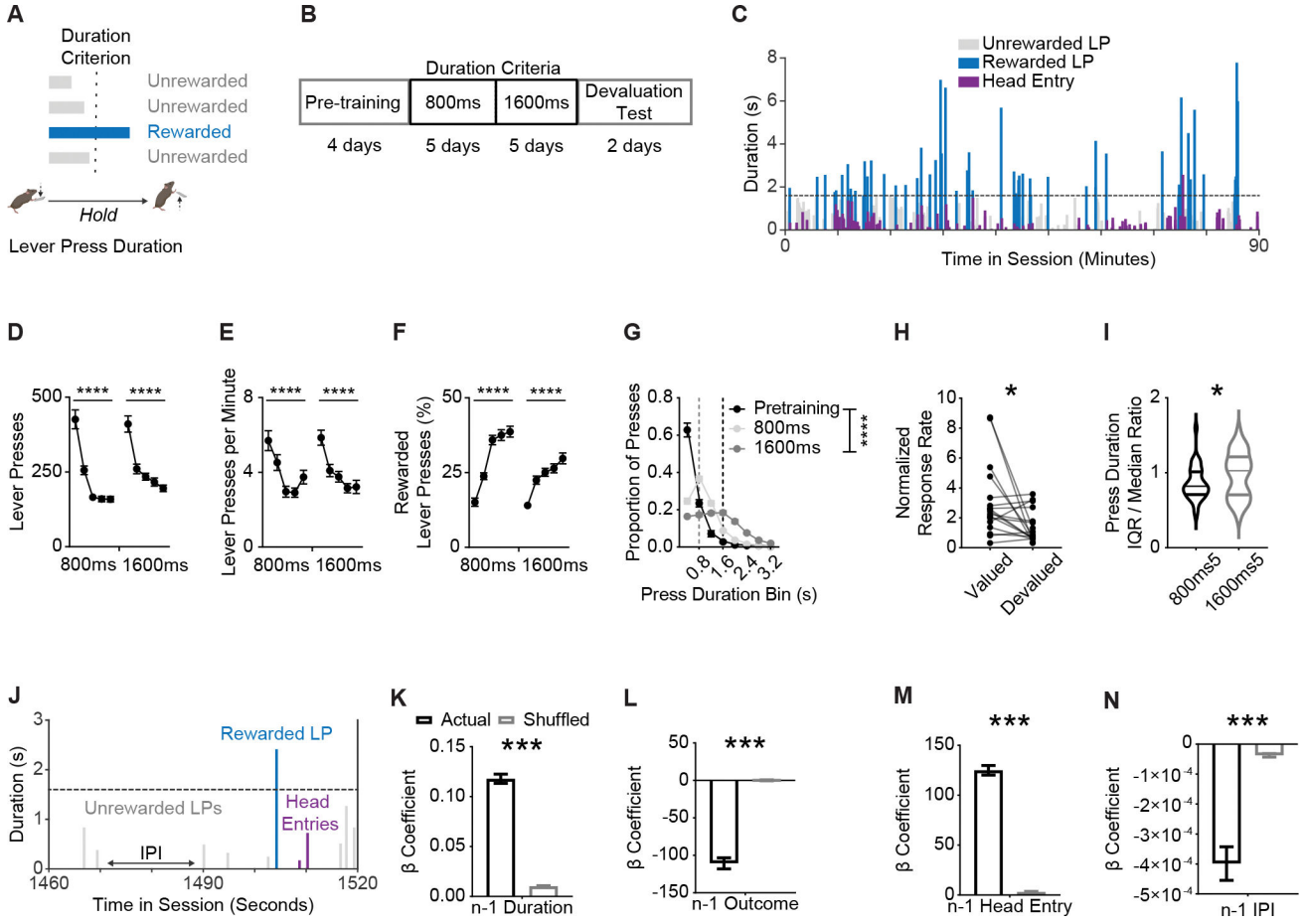


**Figure 1. Mice learned to adjust self-paced, self-generated lever pressing actions across inferred contingency and outcome value changes.**

(A) Behavior schematic demonstrating how mice (C57BL/6J n = 23, 13 males, 10 females; $PV^{cre}$ n = 14, 11 males, 3 females; no effect or interaction of genotype or sex on any behavior measure, thus groups were combined in subsequent analyses) must press and hold down a lever beyond a minimum duration to earn a food reward.

(B) Training schedule for the lever press hold down task. Pretraining sessions were followed by sessions with a minimum duration criterion. Devaluation testing procedures occurred thereafter.

(C) Representative data from one mouse showing variability of lever pressing and head entry behavior within a session. Dashed line indicates 1600 ms criterion.

(D–E) Total lever presses (D), (E) lever pressing rate and (F) percentage of lever presses that exceeded the duration criterion across sessions.

(G) Histogram of lever press durations (400 ms bins) averaged for all pretraining, 800 ms, and 1600 ms duration criterion sessions.

(H) Normalized response rates (lever presses per minute) in valued and devalued states throughout devaluation testing.

(I) Ratio of lever press duration Interquartile Range (IQR) and median during final 800 ms and 1600 ms duration criterion sessions.

(J) Zoomed-in behavior from representative data shown in (C).

(K–N) β coefficients of LME model relating current lever press duration (n) to prior (n − 1) press durations (K), press outcome (i.e. was lever press rewarded) (L), head entry (M), and interpress interval (IPI) (N) for actual and order shuffled data. 800 ms and 1600 ms refer to days where the criterion was >800 ms or >1600 ms.

Data points represent mean ± SEM. Significance markers in k-n indicate comparisons to order shuffled data. Shuffled data are mean ± SEM of 1000 order shuffled β coefficients. *p < 0.05, ***p < 0.001, ****p < 0.0001

See also Figure S1 and Table S1 for more data.

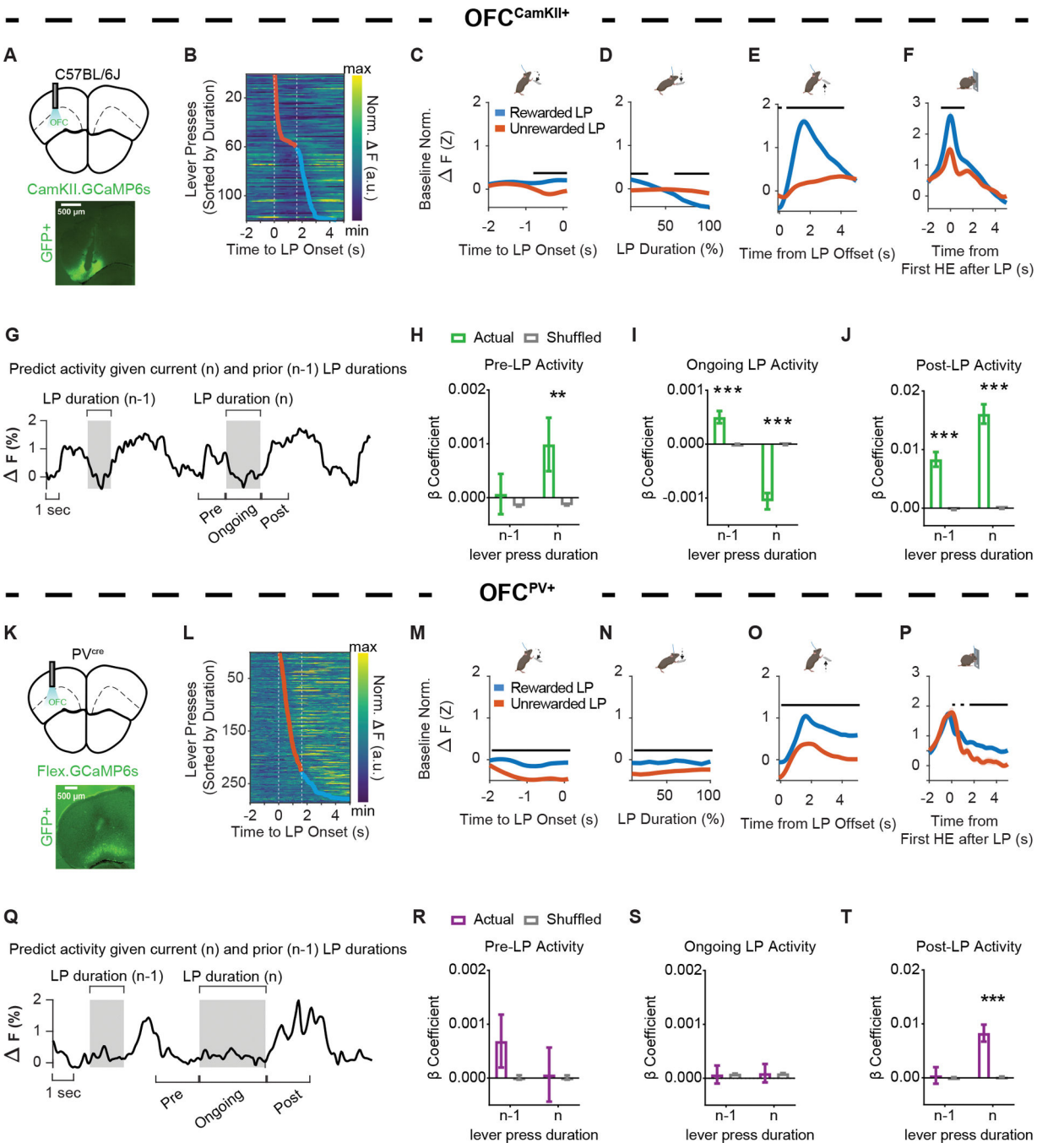


**Figure 2. OFC$^{CamKII+}$, but not OFC$^{PV+}$, $Ca^{2+}$ activity encodes prior action information.**

(A) (top) Anatomical schematic and (bottom) representative histology of (A) OFC$^{CamKII+}$ *in vivo* $Ca^{2+}$ fiber photometry experiments (C57BL/6J, n = 9 mice, 6 males, 3 females). Approximate bregma coordinate AP +2.68 mm.

(B) Representative data heat map of OFC$^{CamKII+}$ normalized fluorescence changes relative to lever press initiation, ordered by lever press duration.

(C–F) $Ca^{2+}$ activity from OFC$^{CamKII+}$ populations z-scored normalized relative to a pre-lever press onset baseline period and aligned to (C) lever press onset, (D) lever press

duration (presented as the relative percentage of total lever press duration), the (E) offset (i.e. termination) of a lever press, and the (F) first head entry made after a lever press.
(G) Representative traces indicating changes in $OFC^{CamKII+}$ $Ca^{2+}$ fluorescence over time.
(H–J) β coefficients from LME models relating $OFC^{CamKII+}$ $Ca^{2+}$ activity to current and prior durations for actual and order shuffled data (H) before press onset (Pre-LP Activity), (I) during the press (Ongoing LP Activity), and (J) after press offset (Post-LP Activity).
(K) (top) Anatomical schematic and (bottom) representative histology of $OFC^{PV+}$ in vivo $Ca^{2+}$ fiber photometry experiments ($PV^{cre}$, n = 8, 5 males, 3 females). Approximate bregma coordinate AP +2.68 mm.
(L) Representative data heat map of $OFC^{PV+}$ normalized fluorescence changes relative to lever press initiation, ordered by lever press duration.
(M–P) $Ca^{2+}$ activity from $OFC^{PV+}$ populations z-scored normalized relative to a pre-lever press onset baseline period and aligned to (M) lever press onset, (N) lever press duration (presented as the relative percentage of total lever press duration), the (O) offset (i.e. termination) of a lever press, and the (P) first head entry made after a lever press.
(Q) Representative traces indicating changes in $OFC^{PV+}$ $Ca^{2+}$ fluorescence over time.
(R–T) β coefficients from LME models relating $OFC^{PV+}$ $Ca^{2+}$ activity to current and prior durations for actual and order shuffled data (R) before press onset (Pre-LP Activity), (S) during the press (Ongoing LP Activity), and (T) after press offset (Post-LP Activity).
LP = lever press, HE = Head Entry. Black lines in C–F and M–P indicate significant differences between Rewarded and Unrewarded lever presses via permutation testing ($p < 0.05$). For B and L, dashed white lines indicate 1600 ms session criterion window. Orange markers indicate the termination of an unrewarded lever press. Blue markers indicate the termination of rewarded lever press. For G and Q, shaded regions indicate the duration of the current (n) or prior (n − 1) lever press. $Ca^{2+}$ activity magnitude predicted by LME models includes −2s to 0s before (Pre), during (Ongoing), or 0 to 2s after (Post) the current lever press.
Shuffled data are the mean ± SEM of 1000 order shuffled β coefficients. ** $p < 0.01$, *** $p < 0.001$.
For a full report of permutation testing p-values related to Figures 2C–F and 2M–P, see publicly available data listed in the key resources table.
See also Figure S2, Table S2 and Table S3 for more data.

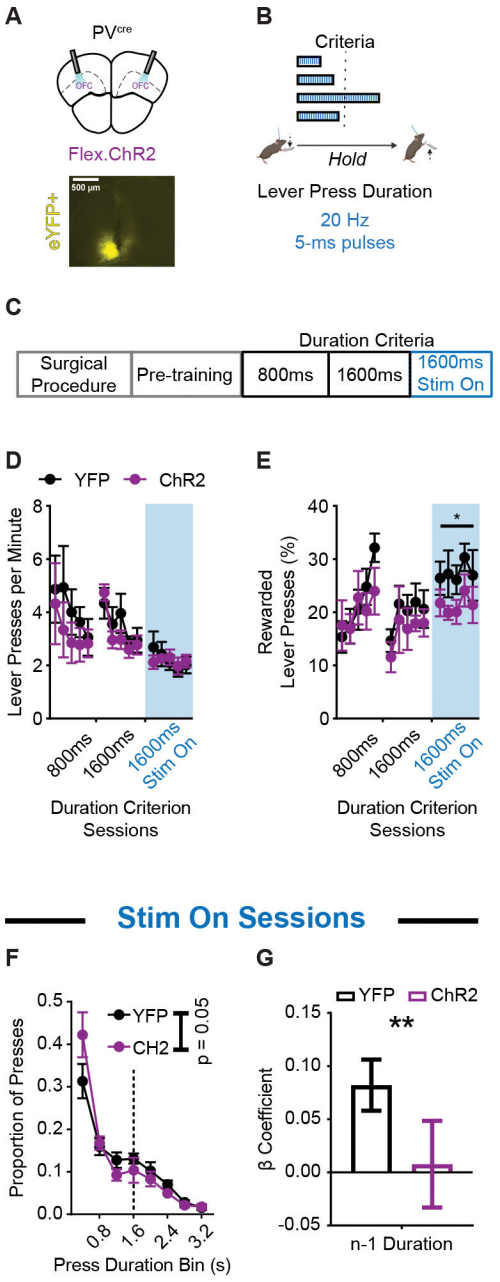


**Figure 3. Optogenetic excitation of OFC$^{PV+}$ populations during action execution reduces rewarded performance and use of prior action information.**

(A) (top) Schematic and (bottom) example histology of ChR2 optogenetic excitation of OFC$^{PV+}$ neurons (PV$^{cre}$ ; n = 6 ChR2, 5 males, 1 female; n = 8 YFP, 5 males, 3 females). Approximate bregma coordinate AP +2.68 mm.

(B) Behavior schematic demonstrating how 470 nm light delivery (20 Hz, 5 ms pulses) occurred for the duration of every lever press made.

(C) Training schedule for optogenetic experiments. Pretraining sessions were followed by sessions with a minimum duration criterion. Sessions in which light was delivered occurred thereafter.

(D and E) Lever pressing rate (D) and (E) percentage of lever presses that exceeded the duration criterion across sessions. Blue shaded region indicates the sessions in which light was delivered. (F) Histogram of lever press durations (400 ms bins) averaged for all 1600 ms duration criterion sessions during which light was delivered.

(G) β coefficients of LME model relating current lever press duration (n) to prior (n − 1) press durations for YFP and ChR2 cohort actual data. Significance marker indicates comparisons to 1000 group shuffled data. 800 ms and 1600 ms refer to days where the criterion was >800 ms or >1600 ms. 1600 ms Stim On refers to days where criterion was >1600 ms and light was delivered.

Data points represent mean ± SEM. *$p < 0.05$, **$p < 0.01$.

See also Table S4 for more data.

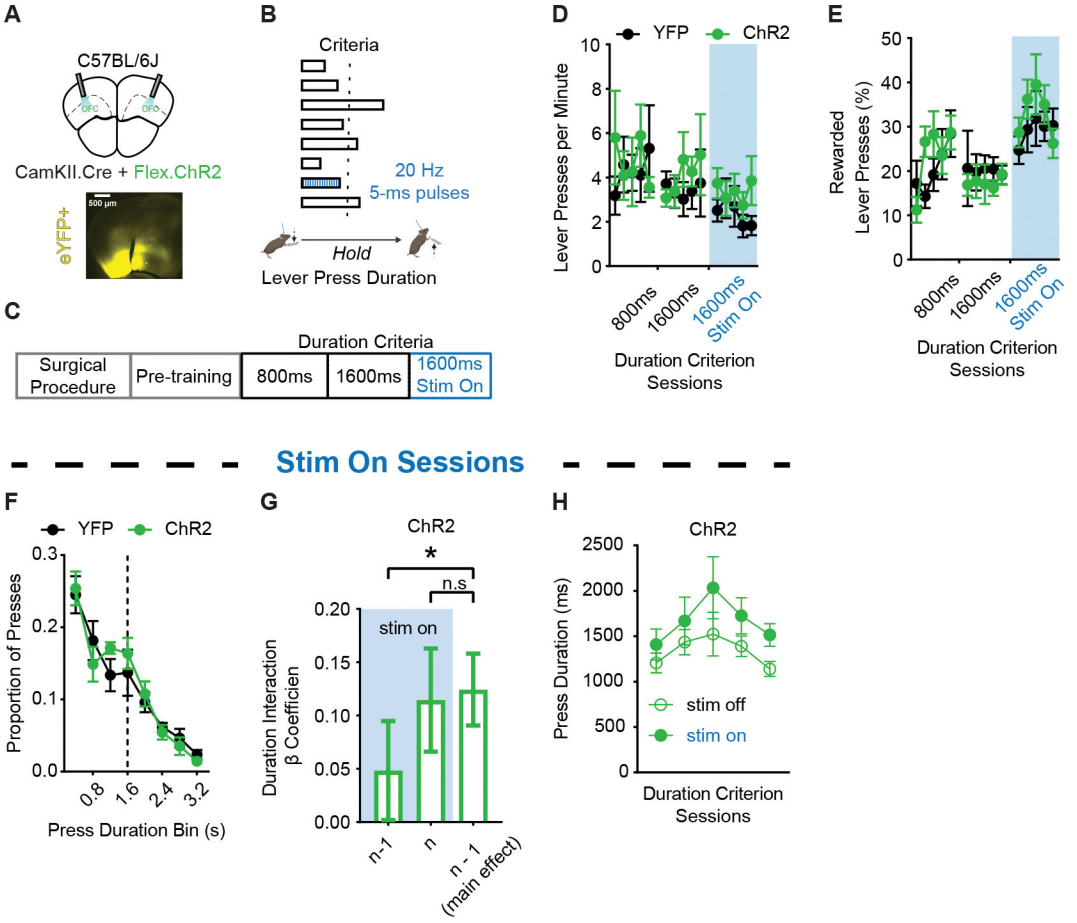


**Figure 4. Selective optogenetic excitation of OFC$^{CamKII+}$ populations during action execution does not impair performance but affects use of prior action information.**

(A) (top) Schematic and (bottom) example histology of ChR2 optogenetic excitation of OFC$^{CamKII+}$ neurons (C57BL/6J; n = 7 ChR2, 5 males, 2 females; n = 5 YFP, 2 males, 3 females). Approximate bregma coordinate AP +2.68 mm.

(B) Behavior schematic demonstrating how 470 nm light delivery (20 Hz, 5 ms pulses) occurred for the duration of every 7th lever press made.

(C) Training schedule for optogenetic experiments. Pretraining sessions were followed by sessions with a minimum duration criterion. Sessions in which light was delivered occurred thereafter.

(D and E) Lever pressing rate (D) and (E) percentage of lever presses that exceeded the duration criterion across sessions. Blue shaded region indicates the sessions in which light was delivered.

(F) Histogram of lever press durations (400 ms bins) averaged for all 1600 ms duration criterion sessions during which light was delivered.

(G) β coefficients of post-hoc LME model relating current lever press duration (n) to prior (n − 1) or current (n) press durations for ChR2 cohort actual data. Significance marker indicates comparisons to 1000 order shuffled data.

(H) Lever press durations of 1600 ms duration criterion sessions during which light was delivered on every 7th lever press, segmented by whether presses were paired with light

activation or not. 800 ms and 1600 ms refer to days where the criterion was >800 ms or >1600 ms. 1600 ms Stim On refers to days where criterion was >1600 ms and light was delivered.

Data points represent mean ± SEM. *$p < 0.05$, *** $p < 0.001$.

See also Figure S3 and Table S5 for more data.

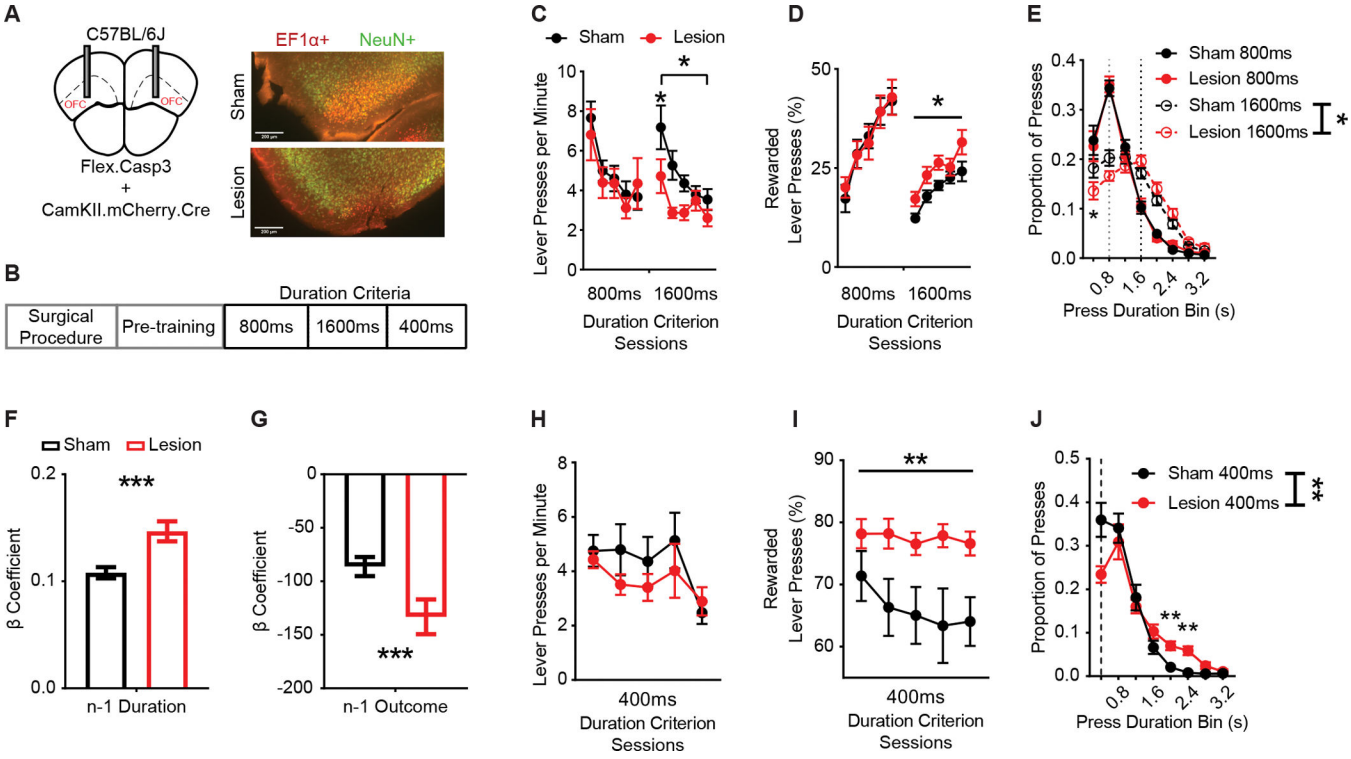


**Figure 5. Pretraining OFC$^{CamKII+}$ lesions increase rewarded performance and use of prior action information.**

(A) (left) Schematic and (right) example histology of sham and Cre-dependent caspase lesion of OFC neurons in mice.(C57BL/6J; Lesion n = 16, 13 males, 3 females; Sham n = 23, 14 males, 9 females). Red indicates AAV-EF1α-DIO-mCherry expression. Green indicates immunohistochemical reactions for neural nuclear protein NeuN. Approximate bregma coordinate AP +2.68 mm.

(B) Training schedule for lesion experiments. Pretraining sessions were followed by sessions with a minimum duration criterion.

(C and D) Lever pressing rate (C) and (D) percentage of lever presses that exceeded the duration criterion across sessions.

(E) Histogram of lever press durations (400 ms bins) averaged for all 800 ms and 1600 ms duration criterion sessions. Dotted lines indicate 800 ms (grey) and 1600 ms (black) duration criteria.

(F and G) β coefficients of LME model relating current lever press duration (n) to prior (n − 1) press durations (F) and press outcome (i.e. was lever press rewarded) (G) for Sham and Lesion cohort actual data. Significance markers indicate comparisons to 1000 group shuffled data.

(H and I) Lever pressing rate (H) and (I) percentage of lever presses that exceeded the 400 ms duration criterion across sessions.

(J) Histogram of lever press durations (400 ms bins) averaged for all 400ms duration criterion sessions. Dotted line indicates 400 ms duration criteria. 400 ms, 800 ms,1600 ms refer to days where the criterion was >400 ms, >800 ms or >1600 ms.

Data points represent mean ± SEM. *p <0 .05, ** p < 0.01, *** p < 0.001.

See also Figure S4 and Table S6 for more data.

## Key Resources Table

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Bacterial and virus strains | | |
| pAAV.CAG.FLEX.GCaMP6s.WPRE.SV40 | Addgene | 100842 |
| rAAVDJ/PAAV-CaMKIIa-GCaMP6s | UNC Vector Core | NA |
| rAAV5/CamKII-hChR2(H134R)-eYFP-WPRE | UNC Vector Core | NA |
| rAAV5/Ef1a-DIO-EYFP | UNC Vector Core | NA |
| rAAV5/Ef1a-DIO-hChR2(H134R)-eYFP | UNC Vector Core | NA |
| rAAV5/Ef1a-DIO-mCherry | UNC Vector Core | NA |
| rAAV5/AAV-Flex-taCasP3-TEVP | UNC Vector Core | NA |
| Chemicals, peptides, and recombinant proteins | | |
| Anti-NeuN (rabbit) Antibody, Alexa Fluor® 488 Conjugate | Sigma-Aldrich | ABN78A4 |
| Experimental models: organisms/strains | | |
| C57BL/6J | The Jackson Laboratory | JAX: 000664 |
| B6.129P2-Pvalbtm1(cre)Arbr/J | The Jackson Laboratory | JAX: 017320 |
| Software and algorithms | | |
| Custom Analysis Code | This paper | https://github.com/gremellab/OFC-Action-Inference |
| GCaMP permutation testing | Jean-Richard-dit-Bressel et al., 2020 | https://doi.org/10.3389/fnmol.2020.00014 |
| Deposited data | | |
| Behavior, photometry, and model output data used in this study | This paper | https://figshare.com/articles/dataset/Orbitofrontal_cortex_populations_are_differentially_recruited_to_support_actions_/20997805 |