Cerebral Cortex, July 2018;28: 2313–2325

doi: 10.1093/cercor/bhx132
Advance Access Publication Date: 25 May 2017
Original Article

ORIGINAL ARTICLE

Insular and Ventrolateral Orbitofrontal Cortices Differentially Contribute to Goal-Directed Behavior in Rodents

Shauna L. Parkes[1,2,3], Pascal M. Ravassard[1,3], Juan-Carlos Cerpa[1,3], Mathieu Wolff[1,3], Guillaume Ferreira[2,3] and Etienne Coutureau[1,3]

[1]CNRS, Institut de Neurosciences Cognitives et Intégratives d'Aquitaine, UMR 5287, 33076 Bordeaux, France, [2]INRA, Nutrition et Neurobiologie Intégrée, UMR 1286, 33076 Bordeaux, France and [3]Universite de Bordeaux, 33076 Bordeaux, France

Address correspondence to Shauna L. Parkes, Université de Bordeaux—site Carreire, Batiment Les Tulipes, Hopital Charles Perrens, 33076 Bordeaux Cedex, France. Email: shauna.parkes@u-bordeaux.fr

Abstract

The medial prefrontal cortex (mPFC) has long been considered a critical site in action control. However, recent evidence indicates that the contribution of cortical areas to goal-directed behavior likely extends beyond mPFC. Here, we examine the function of both insular (IC) and ventrolateral orbitofrontal (vlOFC) cortices in action-dependent learning. We used chemogenetics to study the consequences of IC or vlOFC inhibition on acquisition and performance of instrumental actions using the outcome devaluation task. Rats first learned to associate actions with desirable outcomes. Then, one of these outcomes was devalued and we assessed the rats' choice between the 2 actions. Typically, rats will bias their selection towards the action that delivers the still valued outcome. We show that chemogenetic-induced inhibition of IC during choice abolishes goal-directed control whereas inhibition during instrumental acquisition is without effect. IC is therefore necessary for action selection based on current outcome value. By contrast, vlOFC inhibition during acquisition or the choice test impaired goal-directed behavior but only following a shift in the instrumental contingencies. Our results provide clear evidence that vlOFC plays a critical role in action-dependent learning, which challenges the popular idea that this region of OFC is exclusively involved in stimulus-dependent behaviors.

Key words: action, DREADD, instrumental conditioning, outcome devaluation, reversal

Adaptive decision making requires that organisms track the consequences of their actions. The cornerstone of flexible action selection is up-to-date knowledge regarding both the causal relationship between actions and outcomes and the desirability or value of those outcomes. Such behavior is defined as goal-directed (Dickinson 1985; Balleine and Dickinson 1998; Hilario and Costa 2008; Rangel et al. 2008; Balleine and O'Doherty 2010; Dolan and Dayan 2013; Murray and Rudebeck 2013).

Over the past few decades, the prelimbic region (PL) of the medial prefrontal cortex (mPFC) has been heavily implicated in goal-directed behavior. Specifically, PL is necessary to learn the relationship between actions and their outcomes (Corbit and Balleine 2003; Killcross and Coutureau 2003; Tran-Tu-Yen et al. 2009). However, other cortical regions are now emerging as key players in action selection. The gustatory region of the insular cortex (IC) is required to recall the current value of outcomes to guide choice between competing actions (Parkes and Balleine 2013; Parkes et al. 2015) but, unlike the PL, there is no evidence that IC is involved in action–outcome (A–O) learning per se (Parkes et al. 2016a).

By contrast, the role of ventrolateral orbitofrontal cortex (vlOFC) in instrumental behavior is highly controversial. Current opinion insists that vlOFC is not required for action-guided behavior (Roberts 2006; Ostlund and Balleine 2007a, 2007b; Rudebeck et al. 2008; Balleine et al. 2011; Fellows 2011; Luk and Wallis 2013). Rats with lesions of lateral or ventrolateral OFC and macaque monkeys with lesions of lateral OFC (including areas 11 and 13) are perfectly able to bias their behavior towards an action that delivers a more valuable outcome and to inhibit their responding for an outcome that is rendered less favorable (Ostlund and Balleine 2007b; Rudebeck et al., 2008; Balleine et al. 2011). Yet, more recent reports suggest an involvement of lateral and ventral OFC regions in tasks that also appear to rely on A–O associations (Gremel and Costa 2013; Rhodes and Murray 2013; Gremel et al. 2016; Fiuzat et al. 2017; Zimmermann et al. 2017). These discrepancies may result from differential task requirements.

To investigate this possibility, we assessed the contribution of IC and vlOFC to goal-directed action by manipulating task demands. We used the instrumental outcome devaluation task in which rats learn to associate 2 actions with 2 distinct rewarding outcomes. One of the outcomes is then devalued and the propensity of the subject to perform each action is evaluated. Typically, rats will bias their choice towards the action that delivers the valued outcome. In Experiments 1a and 2a, we adapted the training protocol to include an instrumental pretraining phase, during which both actions earned a common reward, before specific A–O associations were introduced (Corbit and Janak 2010; Corbit et al. 2013; Hart and Balleine 2016). Rats were therefore required to update previously established A–O associations. In Experiments 1b and 2b, the pretraining phase was omitted. We demonstrate that chemogenetic-induced inhibition of IC during action selection abolishes goal-directed control, regardless of training parameters. By contrast, we observed a deficit in goal-directed behavior after inhibition of vlOFC but only when the pretraining phase was included. We hypothesized that this discrepancy reflects the inability of rats with vlOFC inhibition to encode and use the new instrumental contingencies. We therefore assessed the impact of vlOFC inhibition on instrumental reversal learning (Experiment 2b; e.g., Bradfield et al. 2013) and show that vlOFC inhibition indeed hinders goal-directed action control but only following a shift in instrumental contingency.

Materials and Methods

Subjects

Subjects were 110 experimentally naïve, male Long-Evans rats aged 3–4 months (Janvier, France). Rats were housed in pairs in plastic boxes located in a climate controlled room maintained on a 12 h light/dark cycle (lights on at 07:00). All behavior occurred during the light phase of the cycle. Rats were handled daily for 5 days before the behavioral procedures and were put on food restriction 2 days before behavior to maintain them at approximately 90% of their ad libitum feeding weight. Experiments were conducted in agreement with French (council directive 2013-118, 1 February 2013) and international (directive 2010-63, 22 September 2010, European Community) legislations and received approval # 5012053-A from the local Ethics Committee.

Viral Vector

An adeno-associated viral vector carrying the inhibitory hM4Di designer receptor exclusively activated by designer drugs (DREADDs; Armbruster et al. 2007; Rogan and Roth 2011) was obtained from University of North Carolina Vector Core (Chapel Hill, NC). The vector used was AAV8-CaMKII-hM4D(Gi)-mCherry (3–4 × 10^{12} vp/ml). The exogenous ligand, clozapine-N-oxide (CNO; Enzo Life Sciences) was dissolved in 0.9% saline containing 0.5% of dimethyl sulfoxide (Sigma) to obtain a final concentration of 1 mg/ml. CNO was injected intraperitoneally (1 mg/kg) 45 min before behavior.

Surgery

Rats were anaesthetized using Isoflurane (5% induction; 1–2% maintenance) and mounted on a stereotaxic apparatus (Kopf). Rats were subcutaneously injected with 0.05 mg/kg buprenorphine (Buprècare) and the incision site was treated with the local anesthetic xylocaine. The viral vector was infused using repeated pressure pulses delivered via a glass micropipette connected to a picospritzer (Picospritzer III, Parker). For IC, 1 µl of AAV was injected over 5 min at 2 sites in each hemisphere, i.e., 2 µl per hemisphere. The IC co-ordinates were: +0.3 anterior–posterior, ±5.5 medial–lateral, −7.3 dorsal–ventral (from the skull surface) and +1.3 anterior–posterior, ±5.5 medial–lateral, −7.3 dorsal–ventral. For orbitofrontal cortex (OFC), 1 µl of AAV was also injected at 2 sites in each hemisphere to target the ventral and lateral regions. The OFC co-ordinates were: +3.7 anterior–posterior, ±2.0 medial–lateral, −5.0 dorsal–ventral and +3.2 anterior–posterior, ±2.8 medial–lateral, −5.2 dorsal–ventral. All co-ordinates are given in millimeters from bregma (Paxinos and Watson 2014). The glass pipette was left in place for an additional 5 min after infusions to permit diffusion of the virus. Rats were allowed at least 5 weeks to recover before the start of the behavioral procedures, during which time they were monitored daily and weighed. For electrophysiological recordings, the virus was injected bilaterally into OFC at the 2 sites.

Immunohistochemistry

Subsequent to behavioral testing, rats were rapidly and deeply anaesthetized with pentobarbital monosodic and perfused transcardially with 4% paraformaldehyde in 0.1 M phosphate buffer. Brains were removed and postfixed in 4% paraformaldehyde overnight. Subsequently, 50 µm coronal sections were cut using a VT1200S Vibratome (Leica Microsystems). Every fourth section was collected to form a series and immunoreactivity was performed for mCherry. Free-floating sections were prepared by rinsing in 0.1 M phosphate buffered saline (PBS) for 20 min, blocked (1 h, PBS 0.1 M, 0.2% Triton-X, 4% normal goat serum) and placed in 1:1000 rabbit anti-RFP (red fluorescent protein; PM005 CliniSciences) at 4°C for 48 h. Sections were then washed in PBS for 20 min and incubated in 1:200 AffiniPure rhodamine goat anti-rabbit (Jackson Immunoresearch; 111-025-003) diluted in PBS for 2 h at room temperature. Sections were washed for 20 min in PB, mounted, and cover-slipped with Fluoromount-G (SouthernBiotech). Sections were imaged using a Nanozoomer slide scanner (Hamamatsu Photonics) and analyzed with the NDP.view 2 freeware (Hamamatsu Photonics).

Behavioral Apparatus

Instrumental training and testing took place in 8 operant chambers (40 cm width × 30 cm depth × 35 cm height, Imetronic, Pessac, France) individually enclosed in sound and light resistant shells. Each chamber was equipped with 2 pellet dispensers that delivered grain or sugar pellets into a magazine

when activated. The chambers contained 2 retractable levers that could be inserted to the left and right of the magazine. A house light illuminated the chamber. Experimental events were controlled and recorded by a computer located in the room. Devaluation occurred in individual polycarbonate feeding cages located in a different room to the operant chambers.

Behavioral Procedures

We manipulated the training procedures to examine the influence of task demands on subsequent assessment of goal-directed behavior. Each experiment used a between × within design. The between factor was treatment during training (injection of vehicle or CNO) and the within factor was treatment during the choice test (injection of vehicle or CNO). This yielded 4 groups of interest; each labeled according to whether CNO was present (+) or absent (–) during training and test, respectively. At the start of behavioral training, rats in all experiments received 2 sessions of magazine training. During each 40 min session, rats were confined to the operant chamber while either a 20% sucrose solution (0.1 ml; Experiments 1a and 2a) or grain and sugar pellets (45 mg; Experiments 1b and 2b) were delivered, on average, every 60 s. Experiments 1a and 1b examined the effect of IC inhibition and Experiments 2a and 2b examined the effect of vlOFC inhibition.

Pretraining

Rats were initially trained to respond on 2 levers to earn a single, common outcome (20% sucrose solution). During the session, each lever was presented twice for a maximum of 10 min each or until 20 outcomes were earned. The intertrial interval between lever presentations was 2.5 min. The order of the lever presentation was alternated across rats and days. For the first 3 days, lever pressing was continuously reinforced. The probability of the outcome given a response was then shifted using increasing random ratio (RR) schedules: a RR2 schedule was used on Days 4–5, RR3 on Days 6–7 and RR4 on Days 8–9.

Outcome-Specific Instrumental Training

On Days 10–12, 2 distinct rewards were introduced; a grain food pellet (45 mg, BioServ) and a sugar food pellet (45 mg, Test Diet). These sessions were identical to pretraining except that now responding on one lever (e.g., the left lever) delivered one food pellet (e.g., grain) and responding on the other lever (the right lever) delivered the other pellet (sugar) on an RR4 contingency. A–O relationships were counterbalanced across groups. Forty-five minutes before each of these sessions, rats received an intraperitoneal injection of vehicle or CNO.

For experiments without pretraining (Experiments 1b and 2b), rats were immediately trained to perform 2 actions to earn 2 distinct food rewards (e.g., left lever earns a grain pellet and right lever earns a sugar pellet or vice versa). The training procedure was identical to that described above; a fixed-ratio one schedule was used for the first 3 days, a RR2 schedule was used on Days 4–5, RR3 on Days 6–7 and RR4 on Days 8–12. Forty-five minutes before each training session, rats received an injection of vehicle or CNO.

Devaluation Test

Twenty-four hours after the final training session, rats received ad libitum access to one of the 2 outcomes (20 g) for 1 h in familiar feeding cages to induce sensory (not metabolic) satiety (Rolls 1986; Hetherington and Rolls 1996). Immediately after, rats were injected with either vehicle or CNO and, 45 min later, they were given a 10 min choice test in which both levers were available but no outcome was delivered. We have previously demonstrated that satiety-induced devaluation is intact up to 2 h following satiation (Parkes et al. 2016b). The next day, rats were retrained (under vehicle or CNO) and 24 h later they were given a second devaluation test. For the second test, rats that had previously received a vehicle injection now received CNO, whereas rats that had previously received CNO now received vehicle.

Immediately after each instrumental test, rats were returned to the feeding cages and given a consumption test of satiety-induced devaluation. Rats received 10 min access to both food pellets (10 g) and the total amount consumed of each outcome (valued and devalued) was measured. The aim of the consumption test was to ensure that satiety-induced devaluation was effective and that CNO injections did not disrupt the rats' ability to distinguish between the sensory features of the 2 food outcomes.

Reversal of the Instrumental Contingencies

In Experiment 2b, following outcome devaluation testing, the same rats then underwent reversal training such that they were required to learn the reversed instrumental contingencies (e.g., the left lever now earned sugar pellets rather than grain pellets and the right lever now earned grain pellets rather than sugar pellets). Reversal training sessions were identical to outcome-specific instrumental training. Rats received 5 reversal sessions in total on an RR4 schedule of reinforcement and vehicle or CNO was injected 45 min before each session. Rats that received vehicle during initial specific instrumental training also received vehicle during reversal training. Similarly, rats that received CNO during initial training also received CNO during reversal training.

Outcome devaluation tests were conducted after reversal training in the same manner as that previously described. As before, vehicle or CNO was injected immediately after satiation and rats were placed in the operant cages 45 min after the injection. Consumption tests were conducted after each instrumental test.

Electrophysiology

Dissection and Preparation of Acute Brain Slices

Following 6–7 weeks of viral incubation, experimentally naïve rats (n = 15) were deeply anaesthetized with a lethal injection of pentobarbital monosodic (150 mg/kg), and intracardially perfused with oxygenated, ice-cold dissection solution. Brains were then promptly extracted and chilled in the same buffer, and 450 µm-thick coronal slices containing vlOFC were cut using a VT1200S Vibratome. Slices recovered at 35°C for 15 min in a recovery chamber containing the dissection solution. The procedure was adapted from Ting et al. (2014) using (in mM): 93 N-Methyl-D-glucamine (NMDG), 30 $NaHCO_3$, 1.2 NaH_2PO_4, 2.5 KCl, 10 $MgSO_4$, 0.5 $CaCl_2$, 20 HEPES, 25 Glucose, 5 ascorbic acid, 3 pyruvic acid, and 12 N-acetyl-L-cysteine (NAC) equilibrated with 95–5% O_2–CO_2. Slices were then transferred to an incubation chamber at room temperature containing: 92 NaCl, 30 $NaHCO_3$, 1.2 NaH_2PO_4, 2.5 KCl, 2 $MgSO_4$, 2 $CaCl_2$, 20 HEPES, 25 Glucose, 5 ascorbic acid, 3 pyruvic acid, 12 NAC. Following at least 1 h recovery, slices were placed in the recording chamber continuously perfused (2 ml/min) with aCSF containing: 124

NaCl, 24 $NaHCO_3$, 1.2 NaH_2PO_4, 2.5 KCl, 2 $MgSO_4$, 2 $CaCl_2$, 5 HEPES, and 12.5 glucose.

Whole-Cell Recordings

The region of vlOFC was determined based on anatomical landmarks (cortical curvature ventral to the forceps minor of the corpus callosum/claustrum and dorsal to the olfactory nucleus) and neurons from layers 2/3 and 5 were targeted. Transfected and nontransfected neurons were identified under visual guidance using both infrared and epifluorescence illumination with an Eclipse FN1 Nikon microscope equipped with differential interference contrast (DIC). Whole-cell patch-clamp recordings were acquired with Multiclamp 700B amplifier (Molecular Devices) and AxoGraph X software. Data were sampled at 10 kHz and low-pass filtered at 6 kHz using a Bessel filter. Borosilicate pipettes (4–7 MΩ) were filled with internal solution containing (in mM): 120 K-gluconate, 20 KCl, 0.1 $MgCl_2$, 1 EGTA, 10 HEPES, 0.1 $CaCl_2$, 0.1 GTP, 0.2 cAMP, 0.1 Leupeptin, 77 D (–)-Mannitol, 3 Na_2-ATP, pH 7.3 and 300 mOsm/L and 0.1% biocytin was added for labeling sampled neurons. Liquid junction potential of −9 mV was corrected. Cell membrane potential recorded in current-clamp configuration was held at approximately −70 mV and a 250 ms positive current pulse was injected to elicit action potentials (firing rate ≥8 Hz) every 10 s. Following at least 5 min of stable recording, CNO (10 μM) was bath applied for 10 min, and recording continued for an additional 15 min during the wash period. Immediately before, following 10 min CNO application and after 15 min wash, current–voltage (I/V) relationship was tested using 500 ms current pulses of incrementing intensity (20 pA steps, 1 Hz) in order to compute membrane conductance and rheobase. Cells with high initial resting membrane potential (Vm > −30 mV), series resistance (Rs > 40 MΩ) or spike amplitude variation (>50%) were excluded from the analysis ($n = 3$). Putative inhibitory neurons (fast action potentials, high frequency rate, low adaptation) were also excluded ($n = 5$). No significant difference between transfected (hM4Di, $n = 8$) and nontransfected (No-hM4Di, $n = 9$) cells was observed for resting potential ($t_{15} = 0.22$, $P = 0.83$), series resistance (initial: $t_{15} = 0.77$, $P = 0.45$; final: $t_{14} = 1.47$, $P = 0.16$), or holding current values ($t_{15} = 1.20$, $P = 0.25$).

Histology

Slices were processed as described above for mCherry and biocytin was revealed on the same slice by incubation in FITC Streptavidin (1:300, Vector Laboratories) for 2 h. Stained slices were mounted and cover-slipped before visualization using a Leica VM5500B microscope and Exploranova software.

Data Analyses

All behavioral analyses were conducted using a mixed-model ANOVA followed by simple-effects analyses to establish the source of interactions. Statistical significance was set at $P \leq 0.05$. All test data are presented as a percentage of responding during training to control for any baseline differences in responding for grain versus sugar pellets. Electrophysiological data were tested for normality with the Shapiro–Wilk normality test and then analyzed with unpaired *t* tests or Mann–Whitney *U*-test for group comparisons, and paired *t* tests or Wilcoxon signed-rank test for treatment comparisons. Linear regressions significance was tested with a Fisher test and further slope comparisons analyzed with *t* tests.

Results

IC Regulates Action Selection by Retrieving Current Goal Value

In Experiment 1a, we used an instrumental pretraining phase whereas in Experiment 1b the pretraining phase was removed such that rats were not required to update A–O associations.

Experiment 1a

Histology. Figure 1A shows a representative image of viral expression in IC and a schematic of the largest and smallest expression is shown in Figure 1B. Eleven rats were excluded because of unilateral infection or misplacement of the injection. This yielded

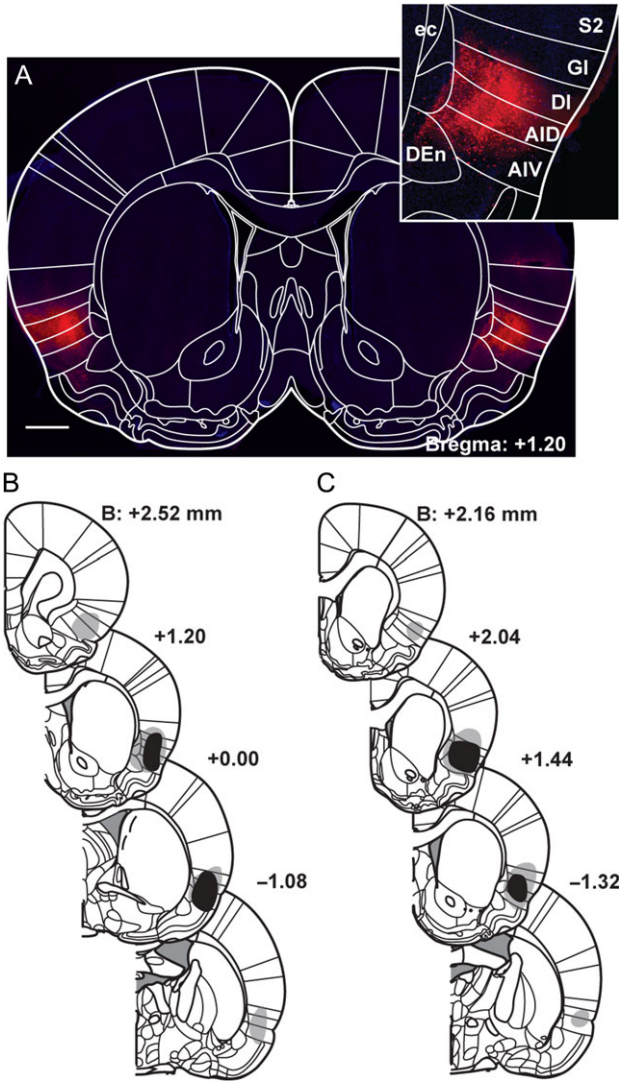

Figure 1. (*A*) Representative photomicrograph illustrating the placement of bilateral AAV8-CaMKII-hM4Di-mCherry injections in insular cortex. Scale bar: 1 mm. Atlas superimposed in white outlines from Paxinos and Watson (seventh edition). Inset: magnification of the area of interest, transfected neurons appear in red. S2: secondary somatosensory cortex, GI: granular insular cortex, DI: dysgranular insular cortex, AID: agranular insular cortex, dorsal, AIV: agranular insular cortex, ventral, ec: external capsule, Den: dorsal endopiriform nucleus. (*B*, *C*) Schematics adapted from Paxinos and Watson (seventh edition) showing the largest (grey) and smallest (black) viral infection for rats included in Experiment 1a (*B*) and 1b (*C*). The vast majority of fluorescence in both experiments was observed between +1.68 and −0.48 mm from bregma.

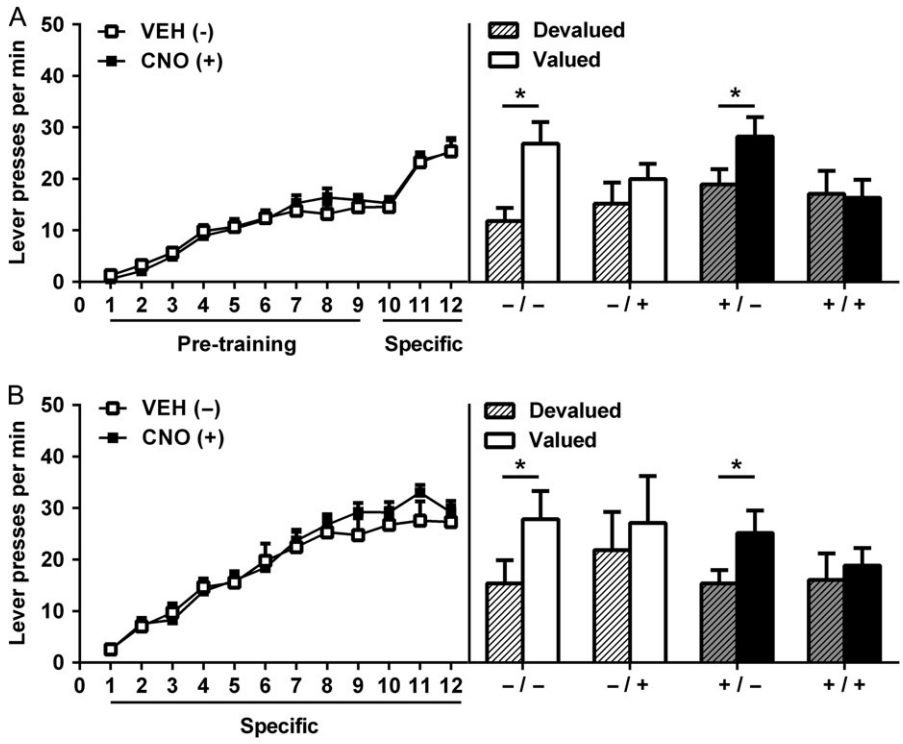

Figure 2. IC regulates action selection by recalling current goal value. (A) Experiment 1a. Left, rate of lever pressing [average +standard error of the mean (SEM)] across pretraining and outcome-specific training averaged across the 2 levers. Rats were injected with either vehicle (white squares) or CNO (black squares) on Days 10–12. Right, average (+SEM) lever presses per minute during the outcome devaluation test. Groups are labeled according to their treatment during training and test, respectively. –: Vehicle injection, +: CNO injection. Groups that received CNO during training are shown in the darker bars. (B) Experiment 1b. Left, rate of lever pressing (average +SEM) across training averaged across the 2 levers. Rats were injected with either vehicle (white squares) or CNO (black squares) on Days 1–12. Right, average (+SEM) lever presses per minute during outcome devaluation test. *Statistically significant difference.

the following between-subject group sizes: vehicle during training, $n = 12$ and CNO during training, $n = 13$.

Instrumental Training. As shown in Figure 2A (left panel), the rate of lever pressing performance increased across pretraining for the common outcome ($F_{1,23} = 158.85$, $P = 0.001$) and did not differ between groups ($F_{1,23} = 0.02$, $P = 0.89$). No day × group interaction was detected ($F_{1,23} = 2.11$, $P = 0.16$). Responding also increased across outcome-specific instrumental training ($F_{1,23} = 53.23$, $P = 0.001$) and there was no significant difference between vehicle- and CNO-treated rats ($F_{1,23} = 0.02$, $P = 0.89$) or a significant day × group interaction ($F_{1,23} = 0.1$, $P = 0.76$). Therefore, inhibition of IC did not interfere with the rats' ability to perform the instrumental response.

Outcome Devaluation Test. Figure 2A (right panel) shows the results of the outcome devaluation test. Inspection of the figure suggests that selective outcome devaluation was impaired when IC was inhibited during test but not when it was inhibited during training. Statistical analyses confirmed this observation; we detected a significant within-subjects effect of devaluation ($F_{1,23} = 11.92$, $P = 0.002$) but not test treatment ($F_{1,23} = 1.43$, $P = 0.24$) and no significant effect of training group ($F_{1,23} = 0.43$, $P = 0.52$). There was a significant overall interaction between lever and test treatment ($F_{1,23} = 9.98$, $P = 0.004$) and statistical analyses restricted to each between-subjects condition revealed a significant interaction between lever and test treatment ($F_{1,23} = 4.86$, $P = 0.04$ and $F_{1,23} = 5.13$, $P = 0.03$ for rats injected with vehicle or CNO during training, respectively).

Simple-effects analyses conducted on these interactions confirmed that rats receiving vehicle at test (group –/– and +/–) responded more for the valued than devalued outcome ($F_{1,23} = 15.73$, $P = 0.001$ and $F_{1,23} = 6.55$, $P = 0.02$, respectively), however, rats that received CNO during test did not (group –/+: [$F_{1,23} = 1.66$, $P = 0.21$] and group +/+: [$F_{1,23} = 0.04$, $P = 0.84$]). Outcome specific devaluation was therefore abolished when IC was inhibited during test. The amount of the outcome consumed during the devaluation period did not differ between groups ($F_{1,23} < 1.07$, $P > 0.3$; data not shown). Importantly, rats in all conditions ate less of the same (devalued) than the different (valued) food outcome during the consumption test (Supplementary Fig. S1), which indicates that devaluation was effective and CNO did not disrupt the ability to distinguish between the sensory features of the 2 food outcomes.

Experiment 1b

Histology. Six rats were excluded because of unilateral infection or misplacement of the injection. This yielded the following group sizes: vehicle during training, $n = 9$ and CNO during training, $n = 11$. A representation of the largest and smallest viral expression is shown in Figure 1C.

Instrumental Training. As shown in Figure 2B (left panel), lever pressing performance increased across training days ($F_{1,18} = 260.86$, $P < 0.001$) and did not differ between groups ($F_{1,18} = 0.21$, $P = 0.65$). No significant interaction was detected ($F_{1,18} = 1.94$, $P = 0.18$).

Outcome Devaluation Test. Figure 2B (right panel) shows the results of the outcome devaluation test. It appears that we replicated the results of the previous experiment; selective outcome devaluation was only attenuated when CNO was given during test. A mixed-model ANOVA found a significant effect of devaluation ($F_{1,18} = 6.9$, $P = 0.02$) but not test treatment ($F_{1,18} < 0.00$, $P > 0.9$) and no significant effect of training group ($F_{1,18} = 0.64$, $P = 0.44$). There was a marginal significant interaction between devaluation and test treatment ($F_{1,18} = 3.3$, $P = 0.09$). Post hoc analyses indicated that rats receiving vehicle at test tended to respond more for the valued than the devalued outcome ($F_{1,18} = 5.23$, $P = 0.04$ and $F_{1,18} = 4.0$, $P = 0.06$ for groups –/– and +/–, respectively) but rats given CNO during test pressed similarly on both levers ($F_{1,18} = 1.17$, $P = 0.29$ and $F_{1,18} = 0.39$, $P = 0.54$ for groups –/+ and +/+, respectively). The amount of the outcome consumed during the devaluation period did not differ between groups ($F_{1,18} < 2.07$, P values ≥ 0.17; data not shown). Again, all groups ate less of the devalued than the valued food outcome during the consumption test (Figure S1).

Taken together, the results of Experiments 1a and 1b reveal that IC is not required for the acquisition of specific A–O associations but is necessary to retrieve outcome representation to guide choice.

vlOFC Tracks Current A–O Associations

As before, in Experiment 2a we used an instrumental pretraining phase whereas in Experiment 2b the pretraining phase was removed such that rats were not required to update the A–O associations.

Experiment 2a

Histology. Figure 3A shows a representative image of viral expression in vlOFC and a schematic of the largest and smallest expression is shown in Figure 3B. Seven rats were excluded because of unilateral infection or misplacement of the injection. This yielded the following group sizes: vehicle during training, $n = 8$ and CNO during training, $n = 9$.

Instrumental Training. Figure 4A (left panel) shows that lever pressing performance increased across pretraining ($F_{1,15} = 156.35$, $P < 0.001$) and did not differ between groups ($F_{1,15} = 0.01$, $P = 0.92$). No interaction was detected ($F_{1,15} = 0.42$, $P = 0.53$). Responding also increased across outcome-specific instrumental training ($F_{1,15} = 260.29$, $P < 0.001$), and there was no significant difference between vehicle- and CNO-treated rats ($F_{1,15} = 0.03$, $P = 0.87$) or a significant interaction ($F_{1,15} = 1.49$, $P = 0.24$). Therefore, vlOFC inhibition did not interfere with the ability of rats to perform the instrumental response.

Outcome Devaluation Test. Figure 4A (right panel) also shows the results of the outcome devaluation test. Selective outcome devaluation was disrupted when CNO was administered during training or during test. We found a significant effect of devaluation ($F_{1,15} = 6.8$, $P = 0.02$) but not test treatment ($F_{1,15} = 2.49$, $P = 0.14$) and no significant effect of training group ($F_{1,15} = 2.65$, $P = 0.12$). There was a trend towards a significant interaction between devaluation and test treatment ($F_{1,15} = 3.33$, $P = 0.09$). Post hoc analyses indicated that rats pressed more for the valued outcome only when vehicle was given both during training and test (i.e., group –/–; $F_{1,15} = 6.19$, $P = 0.03$). When CNO was given during training, test, or both, rats pressed similarly on the levers ($F_{1,15} < 0.98$, $P \geq 0.34$). The amount of the outcome consumed during the devaluation period did not differ between groups (largest $F_{1,15} = 0.69$, $P = 0.42$; data not shown) and all groups consumed less of the devalued than the valued food outcome during the consumption test (Figure S1).

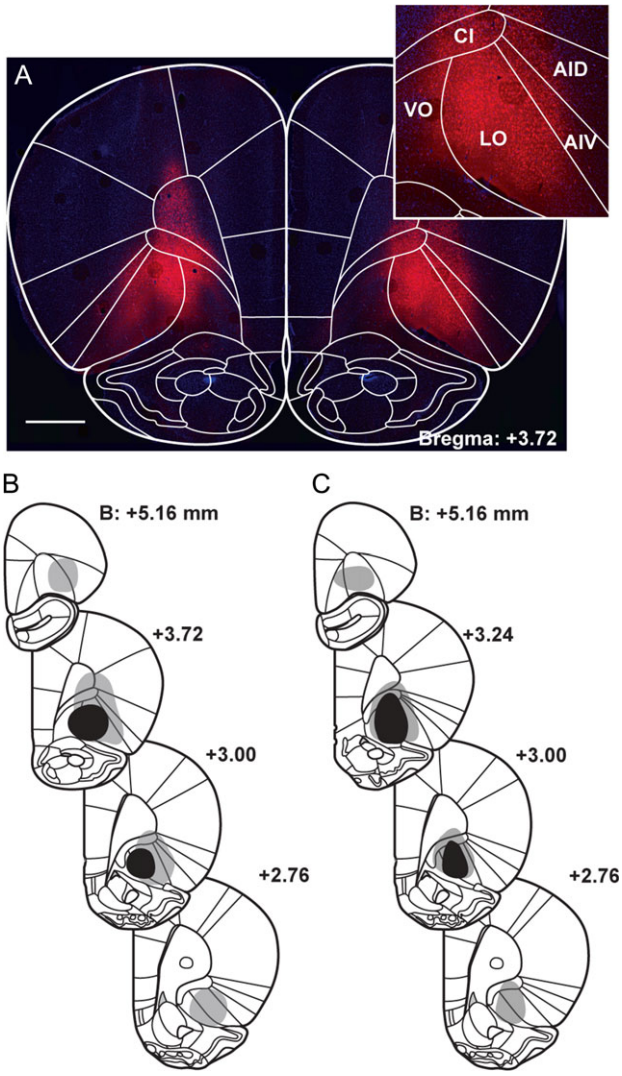

Figure 3. (A) Representative photomicrograph illustrating the placement of bilateral AAV8-CaMKII-hM4Di-mCherry injections in orbitofrontal cortex. Scale bar: 1 mm. Atlas superimposed in white outlines from Paxinos and Watson (seventh edition). Inset: magnification of the area of interest, transfected neurons appear in red. Cl: claustrum, LO: lateral orbital cortex, VO: ventral orbital cortex. (B, C) Schematics adapted from Paxinos and Watson (seventh edition) showing the largest (grey) and smallest (black) viral infection for rats included in Experiment 2a (B) and 2b (C). No viral expression was observed in the medial OFC and limited expression was present in the anterior insular cortex.

Experiment 2b

Histology. Three rats were excluded due to unilateral virus infections. This yielded the following group sizes: vehicle during training, $n = 10$ and CNO during training, $n = 11$. A representation of the largest and smallest viral expression is shown in Figure 3C.

Outcome-Specific Instrumental Training. As illustrated in Figure 4B (left panel), lever pressing performance increased across training ($F_{1,19} = 271.62$, $P < 0.001$) and did not differ between groups

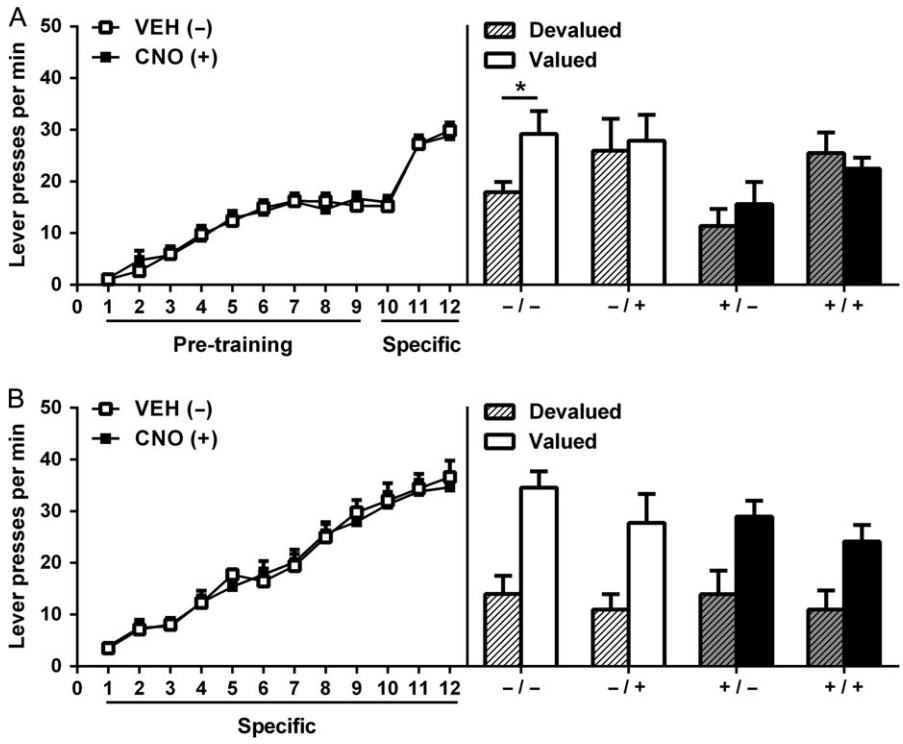

Figure 4. vlOFC tracks current action–outcome associations. (*A*) Experiment 2a. Left, rate of lever pressing (average +SEM) across pretraining and outcome-specific training averaged across the 2 levers. Rats were injected with either vehicle (white squares) or CNO (black squares) on Days 10–12. Right, average (+SEM) lever presses per minute during the outcome devaluation test. Groups are labeled according to their treatment during training and test, respectively. –: Vehicle injection, +: CNO injection. Groups that received CNO during training are shown in the darker bars. (*B*) Experiment 2b. Left, rate of lever pressing (average +SEM) across training averaged across the 2 levers. Rats were injected with either vehicle (white squares) or CNO (black squares) on Days 1–12. Right, average (+SEM) lever presses per minute during the outcome devaluation test. *Statistically significant difference.

($F_{1,19}$ = 0.02, P = 0.89). No significant interaction was detected ($F_{1,19}$ = 0.13, P = 0.72).

Outcome Devaluation Test. Figure 4*B* (right panel) also shows the results of the outcome devaluation test. Inspection of the figure indicates that we did not replicate the results of Experiment 2a; all groups showed selective outcome devaluation, pressing more for the valued than devalued outcome. Indeed, a mixed-model ANOVA found a significant within-subjects effect of devaluation ($F_{1,19}$ = 49.79, P < 0.001) but not test treatment ($F_{1,19}$ = 2.24, P = 0.15) and no significant effect of training group ($F_{1,19}$ = 0.67, P = 0.42). No significant interactions were detected between the 3 factors ($F_{1,19}$ < 0.99, $P \geq$ 0.34). The amount of the outcome consumed during the devaluation period did not differ between groups ($F_{1,19}$ < 1.08, $P \geq$ 0.32) and rats in all conditions ate less of the devalued than the valued food outcome during the consumption test (Figure S1).

Reversal Training. Given that we failed to replicate the results of Experiment 2a, we hypothesized that vlOFC could be responsible for updating changes in instrumental contingencies. We directly tested this prediction by evaluating the ability of these same rats to learn reversed instrumental contingencies and to use this knowledge to guide action selection during vlOFC inhibition. The same rats received 5 sessions of reversal training (Fig. 5, left panel). Lever pressing during reversal training did not differ between vehicle- and CNO-treated rats ($F_{1,19}$ = 0.36, P = 0.56) and there was no significant effect of training day ($F_{1,19}$ = 0.26, P = 0.62) or group × training day interaction ($F_{1,19}$ = 2.8, P = 0.11).

Outcome Devaluation Test Following Reversal. Figure 5 (right panel) shows the results of the reversal test. Inhibition of vlOFC during training or during test impaired selective outcome devaluation based on reversed instrumental contingencies. Statistical analyses found no significant effect of training group ($F_{1,19}$ = 0.1, P = 0.76) or devaluation ($F_{1,19}$ = 2.26, P = 0.15) but there was a significant effect of test treatment ($F_{1,19}$ = 12.32, P = 0.002). A significant interaction between lever and test treatment was also detected ($F_{1,19}$ = 5.19, P = 0.03). Post hoc analyses indicated that rats in the control condition (group –/–) responded more for the valued than the devalued outcome ($F_{1,19}$ = 5.2, P = 0.03). However, rats injected with CNO during training, test or both did not bias their choice towards the lever associated with the valued outcome ($F_{1,19}$ < 0.1, P > 0.5). The amount of the outcome consumed during the devaluation period did not differ between groups (largest $F_{1,19}$ = 1.14, P = 0.3) and all groups rejected the devalued food during the consumption test (Figure S1).

Inhibitory DREADDs Suppressed the Activity of Transfected Cortical Neurons

To validate our chemogenetic approach, we performed in vitro whole-cell patch-clamp recordings in rats injected with AAV8-CaMKII-hM4Di-mCherry in vlOFC. Nontransfected (No-hM4Di) and transfected (hM4Di) cells, identified based on their epifluorescence, were recorded under current-clamp mode. Further immunofluorescence was performed to reveal the co-localization of mCherry-expressing and biocytin-filled neurons (Fig. 6A). Current–voltage (*IV*) relations were conducted by

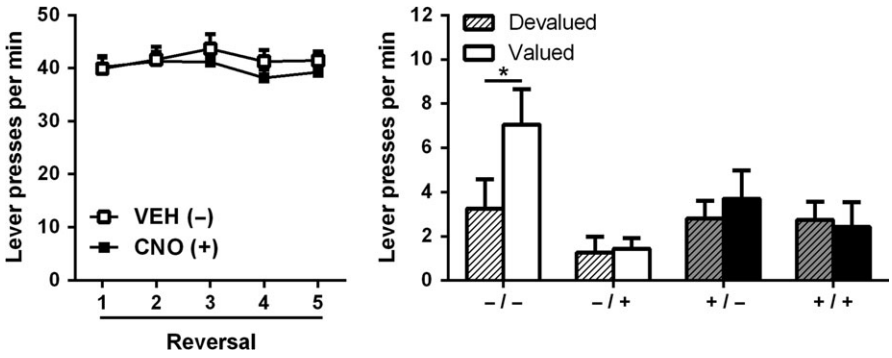

Figure 5. vlOFC and reversal learning. Experiment 2b. Left, rate of lever pressing (average +SEM) across reversal training. Rats were injected with either vehicle (white squares) or CNO (black squares) each day. Right, average (+SEM) lever presses per minute during the outcome devaluation test following reversal training. Devalued and valued refer to the new instrumental contingencies; i.e., the action–outcome associations learnt during reversal training. Groups are labeled according to their treatment during reversal training and test, respectively. –: Vehicle injection, +: CNO injection. Groups that received CNO during training are shown in the darker bars. *Statistically significant difference.

injecting incremental currents in the cells both pretreatment and following 10 min CNO bath application (10 μM), and showed higher membrane conductance and rheobase (supplemental Figure S2) in transfected cells, consistent with the expected opening of G protein-coupled inwardly rectifying potassium channels (GIRKs) via the CNO-induced activation of hM4Di receptors.

The dynamic changes in membrane potentials (Fig. 6B) and firing rates (Fig. 6C) were also measured in transfected and non-transfected cells with continuous recordings. As illustrated, CNO application resulted in a marked membrane hyperpolarization in hM4Di cells (from −68.7 ± 8.2 to −73.8 ± 8.8 mV, paired t-test $t_7 = 4.43$, $P = 0.003$), but not in No-hM4Di cells (from −70.1 ± 8.3 to −70.6 ± 8.3 mV, paired *t*-test $t_8 = 0.65$, $P = 0.53$), and the hyperpolarization was significantly stronger in hM4Di cells (group comparison: $t_{15} = 3.49$, $P = 0.003$). This effect was accompanied by a 60% reduction of the baseline firing rate in the transfected cells (from 17.3 ± 2.1 to 6.9 ± 1.2 Hz, paired *t*-test $t_7 = 5.37$, $P = 0.001$), whereas no significant change was found in No-hM4Di cells (from 15.0 ± 1.8 to 14.6 ± 1.8 Hz, paired *t*-test $t_7 = 0.75$, $P = 0.32$, group comparison: $t_{15} = 2.49$, $P = 0.025$). This result is consistent with a strong silencing of the excitatory neurons activity (−10.5 ± 1.4 Hz for hM4Di cells vs. −0.4 ± 0.5 Hz for No-hM4Di cells, $t_{15} = 4.35$, $P = 0.001$) and was observed in most sampled vlOFC transfected cells (7/8).

Discussion

The present results show that chemogenetic-induced inhibition of IC impaired the performance but not acquisition of goal-directed behavior. Importantly, the latter finding held true regardless of whether the training protocol demanded an updating of instrumental contingencies or not. We also uncovered a novel role for OFC in goal-directed action. Rats with vlOFC inhibition can initially learn and express A–O learning but are impaired when the contingencies are altered. Our electrophysiological results confirmed that CNO application reduced activity by 60% in cortical neurons infected with the inhibitory DREADD, in accordance with previous studies (Bradfield et al. 2015).

Cortical Contributions to Goal-Directed Behavior

In a series of recent studies, we provided evidence that IC is critical for the performance (Parkes and Balleine 2013; Parkes et al. 2015), but not acquisition (Parkes et al. 2016a), of goal-directed behavior. Here, we confirm these findings and show that IC remains uninvolved in acquisition whether rats simply learned to associate 2 actions with 2 different outcomes (Fig. 2B) or were required to update previously established A–O associations that had been acquired during a pretraining phase (Fig. 2A). Our results are also consistent with a previous report that IC lesions do not affect instrumental contingency degradation. Rats with IC lesions decrease their performance of an action when the contingency between that action and its specific outcome is weakened but continue to perform another action for which the A–O contingency is intact (Balleine and Dickinson 2000). To achieve this, rats must have knowledge about the specific A–O associations. Together, these results show that compromised IC function renders rats unable to select an action based on the changing value of its consequences but does not affect the ability to learn and update instrumental contingencies.

By contrast, evidence indicates that the PL region of mPFC is necessary to learn and consolidate instrumental contingencies. Sensitivity to outcome devaluation is abolished in rats with pretraining lesions of PL (Corbit and Balleine 2003; Killcross and Coutureau 2003) or when PL is inactivated during training but not when it is inactivated during test (Tran-Tu-Yen et al. 2009). Moreover, inhibition of mitogen-activated protein kinase/extracellular signal-related kinase in the PL region immediately after acquisition also impairs performance on a subsequent outcome devaluation task (Hart and Balleine 2016). As such, the PL and IC appear to play complementary roles in goal-directed actions; the former being necessary for acquisition but not performance and the latter for the performance but not the acquisition. At present, there is no evidence, of which we are aware, that implicates IC or PL in the incentive learning processes that occur during selective satiation. Instead, this process appears to rely on subcortical regions, namely the basolateral amygdala (Wassum et al. 2009; West et al. 2012; Parkes and Balleine 2013).

The gustatory portion of IC has traditionally been investigated for its role in taste processing (Yamamoto 1984; Maffei et al. 2012; Bermudez-Rattoni 2014). As such, one might expect that inhibition of IC immediately following satiation could interfere with consolidation of this incentive learning. Similar manipulations have indeed been shown to interfere with consolidation of conditioned taste aversion (Gutierrez et al. 1999; Ferreira et al. 2002). However, we believe that this is unlikely given that IC inhibition did not affect performance on the consumption test of satiety-induced devaluation (supplemental

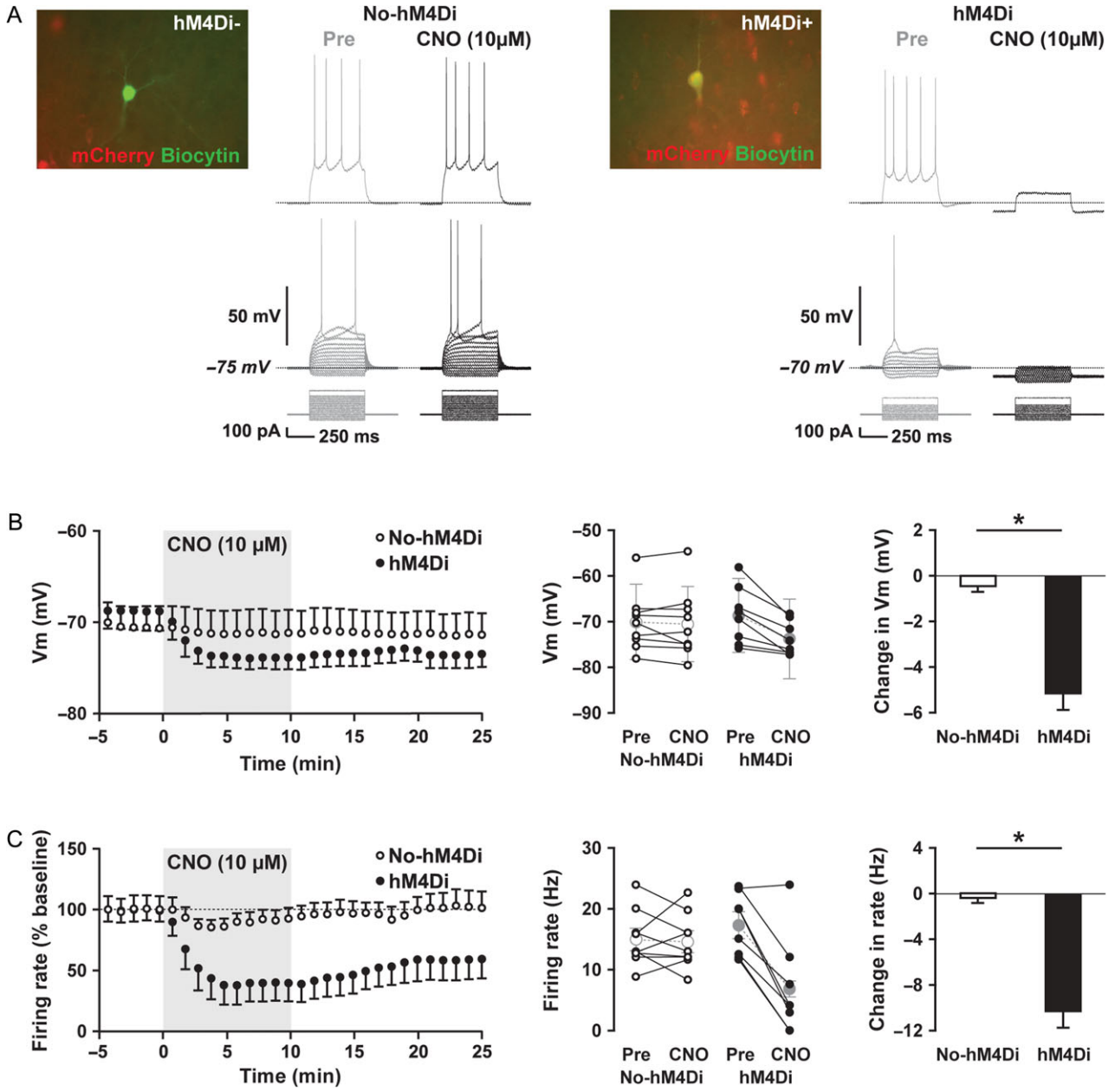

Figure 6. Inhibitory DREADDs suppressed the activity of transfected vlOFC neurons. (*A*) Representative whole-cell current-clamp experiments performed on nontransfected (No-hM4Di, left panel) and transfected neurons (hM4Di, right panel) for which current–voltage relation have been tested by applying incremental current pulses (500 ms duration) prior (condition Pre, grey traces) and after 10 min CNO bath application (10 μM, condition CNO, black traces). Most depolarized voltage trace in response to the highest current intensity appears above the other traces for clarity. Note the absence of action potentials and the lower membrane potential in hM4Di neuron in presence of CNO. Dashed lines represent the membrane potential at holding current (invariant throughout protocol). Insets: immunofluorescent labeling of transfected neurons (mCherry, red) and recorded neurons (biocytin, green). (*B*) Left, continuous whole-cell current-clamp recording showing the time course of the membrane potential prior, during CNO bath application (shaded area), and during wash period for nontransfected (n = 9 No-hM4Di, open circles) and transfected neurons (n = 8 hM4Di, black circles). Middle, individual membrane potential values for each sampled neurons during Pre and CNO conditions, average ± SEM represented in grey outlines. Right, average change in absolute membrane potential showing a significant hyperpolarization for hM4Di compared with nontransfected neurons. (*C*) Left, continuous whole-cell current-clamp recording showing the time course of the firing rate, during CNO bath application (shaded area), and during wash period for nontransfected (n = 9 No-hM4Di, open circles) and transfected neurons (n = 8 hM4Di, black circles). Values are expressed in percentage ± SEM of the baseline period. Middle, individual values of firing rate for each sampled neurons during Pre and CNO conditions, average ± SEM represented in grey outlines. Right, average change in absolute firing rate showing a significant decrease for hM4Di compared with nontransfected neurons. *Statistically significant difference.

Figure S1). Recent evidence has extended the role of IC from taste encoding to encoding stimuli that predict taste or food outcomes (Samuelsen et al. 2012; Tang et al. 2012; Gardner and Fontanini 2014; Kusumoto-Yoshida et al. 2015). Fontanini and colleagues demonstrated that IC neurons differentially respond to stimuli predicting appetitive versus aversive tastes in a go/no-go task (Samuelsen et al. 2012; Gardner and Fontanini 2014) and similar findings were also recently reported in mice engaged in pavlovian conditioning (Kusumoto-Yoshida et al. 2015). Given the present results, it appears that IC plays a

general role in guiding behavior based on the current value of expected food outcomes. This is likely achieved via projections from IC to the ventral striatum (Parkes et al. 2015).

We observed a different pattern of results for vlOFC. Specifically, inhibition of vlOFC disrupted selective outcome devaluation when a pretraining phase was included (Experiment 2a, Fig. 4A) but this disruption was not observed when the pretraining phase was omitted (Experiment 2b, Fig. 4B). We hypothesized that this discrepancy might be attributed to the change in A–O associations that occurs when a pretraining phase is included. To test this hypothesis, we assessed the impact of vlOFC inhibition on reversal learning and showed that vlOFC inhibition indeed impairs the ability of rats to learn and use reversed instrumental contingencies (Experiment 2b, Fig. 5).

Importantly, goal-directed control was attenuated when vlOFC was inhibited during training or during test following a shift in A–O contingencies. Therefore, vlOFC is required to both encode and recall the identity of the expected outcome, particularly when that identity has changed. This result highlights a point of difference in the involvement of vlOFC and PL in goal-directed behavior. PL plays a transient role; it is important for early acquisition but not for long-term storage of the A–O contingencies or for performance (Ostlund and Balleine 2005; Tran-Tu-Yen et al. 2009). PL may therefore encode initial contingency-related information (i.e, encoding the basic causal relationships), which is subsequently stored in posterior dorsomedial striatum (Hart and Balleine 2016). By contrast, vlOFC encodes and stores representations of the specific features of outcomes that are currently associated with responses (or stimuli) and permits the formation of distinct task states to represent updated A–O contingencies (Wilson et al. 2014).

Notably, the impairment we observed following reversal of the instrumental associations was not due to perseveration on the old contingencies and, therefore, goal-directed control may have been restored with additional reversal training (Schoenbaum et al. 2003). As such, vlOFC inhibition does not appear to induce a general learning deficit but rather the inability to immediately recognize changes in the predictive relationship between discrete actions and their specific consequences due to interference with the original contingencies (Gershman and Niv 2012; Bradfield et al. 2013; Stalnaker et al. 2016). It should be noted that the amount of outcome specific training differed across experiments. Indeed, vlOFC inhibition produced a deficit in outcome specific devaluation when the number of specific training sessions was reduced. Perhaps we would not have observed impaired goal-directed control if the number of training sessions was increased. However, others have reported that, regardless of amount of training or levels of difficulty during training, OFC lesions leave acquisition of odor-outcome associations intact but impair reversal learning (Kim and Ragozzino 2005).

Our results are consistent with previous reports that lesions of lateral or ventrolateral OFC leave instrumental outcome specific devaluation intact when the A–O contingencies remain stable (Ostlund and Balleine 2007a, 2007b; Balleine et al. 2011). However, conflicting results have recently been reported in both rodents and primates (Gremel and Costa 2013; Rhodes and Murray 2013; Gremel et al. 2016; Fiuzat et al. 2017; Zimmermann et al. 2017) although, it should be noted that there is some concern as to whether rodent OFC (composed exclusively of agranular cortical areas) is homologous to the larger, granular primate OFC (Preuss 1995). Nevertheless, impaired instrumental outcome specific devaluation has been observed in both mice with compromised function of ventral and lateral OFC (Gremel and Costa 2013) and rhesus monkeys with lesions of the entire OFC, including areas 11, 13, and 14 (Rhodes and Murray 2013; Fiuzat et al. 2017). Importantly, these studies included a substantial pavlovian component; the former using contextual cues to modulate responding and the latter primate studies required pavlovian pretraining, which makes the influence of discrete stimuli difficult to ascertain. Indeed, current opinion largely insists that vlOFC is required for stimulus- (Gallagher et al. 1999; Pickens et al. 2003, 2005; Izquierdo and Murray 2004, 2010; Izquierdo et al. 2004; Machado and Bachevalier 2007; Ostlund and Balleine 2007b; West et al. 2011) but not action-guided behavior (Ostlund and Balleine 2007a, 2007b; Rudebeck et al. 2008; Balleine et al. 2011; Luk and Wallis 2013). Moreover, it was recently proposed that it is the medial region of OFC, not vlOFC, which regulates instrumental actions (Bradfield et al. 2015; Gourley et al. 2016).

An Inclusive Role for vlOFC in Adaptive Behavior

By contrast, our results suggest a general role for vlOFC when expectations are violated, regardless of the nature of the task. This is consistent with the view that behaviors relying on the explicit use of learned associations may be OFC-independent (Schoenbaum and Roesch 2005). Indeed, vlOFC is required for other tasks that generate ambiguity, including pavlovian reversal learning (Jones and Mishkin 1972; Dias et al. 1997; Schoenbaum et al. 2002; Chudasama and Robbins 2003; Rudebeck and Murray 2008), contingency degradation (Ostlund and Balleine 2007b; Alcaraz et al. 2015) and choice behavior guided by learned taste aversion (Ramirez-Lugo et al. 2016). However, it must be noted that vlOFC inhibition causes a deficit in pavlovian devaluation tasks even when the stimulus-outcome contingencies remain unchanged (Pickens et al. 2003, 2005; Izquierdo and Murray 2004; Izquierdo et al. 2004). The reason for this distinction is not immediately clear but perhaps the pavlovian tasks somehow evoke a greater sense of ambiguity; for example, via stimulus generalization.

It has been proposed that OFC (particularly the lateral region) is critical for forming and recognizing task states; representations of current task conditions and beliefs (Wilson et al. 2014). Notably, OFC is especially implicated when task states change without explicit notice (Wilson et al. 2014), as is the case in the present study. Here, inhibition of vlOFC may result in the encoding (inhibition during training) or retrieval (inhibition during test) of a weaker representation of outcome features that, in turn, render the rats unable to form a distinct task state to represent the new A–O contingencies. This failure to form different states may explain why we sometimes observed a general rather than outcome specific devaluation effect. Importantly, the vast majority of the evidence supporting the "state theory" is derived from behavioral tasks with significant pavlovian components. Here, we show that vlOFC tracks current states regardless of the underlying associative structure of the task.

Neural Circuitry of Goal-Directed Behavior

We have demonstrated distinct roles of rodent IC and vlOFC in goal-directed behavior. How different regions of the cortex might interact to support this behavior is currently unknown. The mPFC is reciprocally connected to both vlOFC (Vertes 2004; Hoover and Vertes 2011) and IC (Conde et al. 1995; Shi and Cassell 1998; Gabbott et al. 2003) but it remains to be determined if these areas communicate directly, or indirectly via

transthalamic routes (Sherman 2016), to regulate action selection. Moreover, given the strong intrainsular connections (Fujita et al. 2010) and the direct reciprocal projections between OFC and IC (Shi and Cassell 1998; Fujita et al. 2010; Hoover and Vertes 2011), it is worthwhile to consider that communication between vlOFC and IC is required for integrating current A–O contingencies with goal value to guide adaptive behavior.

Additional research is also required to determine how insular and orbitofrontal cortices make contact with the broader neural circuity supporting goal-directed behavior. Previous evidence suggests that action-guided choice may rely on communication between BLA and IC (Parkes and Balleine 2013) as well as between BLA and OFC (Zeeb and Winstanley 2013). Moreover, pavlovian tasks indicate that BLA and vlOFC function in cooperation to track expected rewards (Schoenbaum et al. 2003; Saddoris et al. 2005; Hampton et al. 2007; Rudebeck et al. 2013; Lucantonio et al. 2015). Given BLA's role in updating and encoding changes in outcome value during satiation (West et al. 2012; Parkes and Balleine 2013), further investigation on the direct involvement of BLA–IC and BLA–vlOFC connections in goal-directed action would be highly beneficial to our current understanding of the neural circuit mediating choice.

Conclusion

Our results challenge the notion that the involvement of vlOFC in outcome processing is restricted to stimulus-guided behavior. We provide clear evidence that vlOFC is recruited to learn and use A–O associations whereas IC plays a more general role in the retrieval of outcome value to guide choice behavior, regardless of contingency changes. These findings not only increase our understanding of the cortical bases of goal-directed action but will also reinforce and advance current theories on the function of OFC in decision making.

Supplementary Material

Supplementary material is available at *Cerebral Cortex* online.

Funding

French National Research Agency (grant number ANR-14-CE13-0014 GOAL to E.C. and G.F.), the French National Institute for Agricultural Research (INRA) and AgreenSkills (grant agreement number FP7-267196 to S.L.P.), which has received funding from the EU's Seventh Framework Programme.

Notes

The microscopy was completed at the Bordeaux Imaging Center, a service unit of CNRS-INSERM and Bordeaux University and member of the national infrastructure, France BioImaging. Electrophysiological recordings were conducted at the PhyCell platform (INCIA UMR 5287, Bordeaux University). We thank S. Bertrand and A.E. Allain for their technical and theoretical support in electrophysiology experiments and Y. Salafranque for the care provided to the rats during experiments. *Conflict of Interest*: None declared.

References

Alcaraz F, Marchand AR, Vidal E, Guillou A, Faugere A, Coutureau E, Wolff M. 2015. Flexible use of predictive cues beyond the orbitofrontal cortex: role of the submedius thalamic nucleus. J Neurosci. 35:13183–13193.

Armbruster BN, Li X, Pausch MH, Herlitze S, Roth BL. 2007. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. Proc Natl Acad Sci USA. 104:5163–5168.

Balleine BW, Dickinson A. 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology. 37:407–419.

Balleine BW, Dickinson A. 2000. The effect of lesions of the insular cortex on instrumental conditioning: evidence for a role in incentive memory. J Neurosci. 20:8954–8964.

Balleine BW, Leung BK, Ostlund SB. 2011. The orbitofrontal cortex, predicted value, and choice. Ann N Y Acad Sci. 1239:43–50.

Balleine BW, O'Doherty JP. 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology. 35:48–69.

Bermudez-Rattoni F. 2014. The forgotten insular cortex: its role on recognition memory formation. Neurobiol Learn Mem. 109:207–216.

Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW. 2013. The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. Neuron. 79:153–166.

Bradfield LA, Dezfouli A, Van Holstein M, Chieng B, Balleine BW. 2015. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. Neuron. 88:1268–1280.

Chudasama Y, Robbins TW. 2003. Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. J Neurosci. 23:8771–8780.

Conde F, Maire-Lepoivre E, Audinat E, Crepel F. 1995. Afferent connections of the medial frontal cortex of the rat. II. Cortical and subcortical afferents. J Comp Neurol. 352:567–593.

Corbit LH, Balleine BW. 2003. The role of prelimbic cortex in instrumental conditioning. Behav Brain Res. 146:145–157.

Corbit LH, Janak PH. 2010. Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. Eur J Neurosci. 31:1312–1321.

Corbit LH, Leung BK, Balleine BW. 2013. The role of the amygdala-striatal pathway in the acquisition and performance of goal-directed instrumental actions. J Neurosci. 33:17682–17690.

Dias R, Robbins TW, Roberts AC. 1997. Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: restriction to novel situations and independence from "on-line" processing. J Neurosci. 17:9285–9297.

Dickinson A. 1985. Actions and habits: the development of behavioral autonomy. Philos Trans R Soc Lond B Biol Sci. 308:67–78.

Dolan RJ, Dayan P. 2013. Goals and habits in the brain. Neuron. 80:312–325.

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.

Ferreira G, Gutierrez R, De La Cruz V, Bermudez-Rattoni F. 2002. Differential involvement of cortical muscarinic and NMDA receptors in short- and long-term taste aversion memory. Eur J Neurosci. 16:1139–1145.

Fiuzat EC, Rhodes SE, Murray EA. 2017. The role of orbitofrontal-amygdala interactions in updating action-outcome valuations in macaques. J Neurosci. 37:2463–2470.

Fujita S, Adachi K, Koshikawa N, Kobayashi M. 2010. Spatiotemporal dynamics of excitation in rat insular cortex: intrinsic corticocortical circuit regulates caudal-rostro excitatory propagation from the insular to frontal cortex. Neuroscience. 165:278–292.

Gabbott PL, Warner TA, Jays PR, Bacon SJ. 2003. Areal and synaptic interconnectivity of prelimbic (area 32), infralimbic (area 25) and insular cortices in the rat. Brain Res. 993:59–71.

Gallagher M, Mcmahan RW, Schoenbaum G. 1999. Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci. 19:6610–6614.

Gardner MP, Fontanini A. 2014. Encoding and tracking of outcome-specific expectancy in the gustatory cortex of alert rats. J Neurosci. 34:13000–13017.

Gershman SJ, Niv Y. 2012. Exploring a latent cause theory of classical conditioning. Learn Behav. 40:255–268.

Gourley SL, Zimmermann KS, Allen AG, Taylor JR. 2016. The medial orbitofrontal cortex regulates sensitivity to outcome value. J Neurosci. 36:4600–4613.

Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, Deisseroth K, Lovinger DM, Costa RM. 2016. Endocannabinoid modulation of orbitostriatal circuits gates habit formation. Neuron. 90:1312–1324.

Gremel CM, Costa RM. 2013. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat Commun. 4:2264.

Gutierrez H, Hernandez-Echeagaray E, Ramirez-Amaya V, Bermudez-Rattoni F. 1999. Blockade of N-methyl-D-aspartate receptors in the insular cortex disrupts taste aversion and spatial memory formation. Neuroscience. 89:751–758.

Hampton AN, Adolphs R, Tyszka MJ, O'doherty JP. 2007. Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. Neuron. 55: 545–555.

Hart G, Balleine BW. 2016. Consolidation of goal-directed action depends on MAPK/ERK signaling in rodent prelimbic cortex. J Neurosci. 36:11974–11986.

Hetherington MM, Rolls BJ. 1996. Sensory-specific satiety: theoretical frameworks and central characteristics. In: Capaldi ED, editor. Why we eat what we eat: the psychology of eating. Washington, DC, USA: American Psychological Association.

Hilario MR, Costa RM. 2008. High on habits. Front Neurosci. 2: 208–217.

Hoover WB, Vertes RP. 2011. Projections of the medial orbital and ventral orbital cortex in the rat. J Comp Neurol. 519: 3766–3801.

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

Izquierdo A, Murray EA. 2010. Functional interaction of medial mediodorsal thalamic nucleus but not nucleus accumbens with amygdala and orbital prefrontal cortex is essential for adaptive response selection after reinforcer devaluation. J Neurosci. 30:661–669.

Izquierdo A, Suda RK, 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.

Jones B, Mishkin M. 1972. Limbic lesions and the problem of stimulus—reinforcement associations. Exp Neurol. 36: 362–377.

Killcross S, Coutureau E. 2003. Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb Cortex. 13:400–408.

Kim J, Ragozzino ME. 2005. The involvement of the orbitofrontal cortex in learning under changing task contingencies. Neurobiol Learn Mem. 83:125–133.

Kusumoto-Yoshida I, Liu H, Chen BT, Fontanini A, Bonci A. 2015. Central role for the insular cortex in mediating conditioned responses to anticipatory cues. Proc Natl Acad Sci USA. 112:1190–1195.

Lucantonio F, Gardner MP, Mirenzi A, Newman LE, Takahashi YK, Schoenbaum G. 2015. Neural estimates of imagined outcomes in basolateral amygdala depend on orbitofrontal cortex. J Neurosci. 35:16521–16530.

Luk CH, Wallis JD. 2013. Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. J Neurosci. 33:1864–1871.

Machado CJ, Bachevalier J. 2007. The effects of selective amygdala, orbital frontal cortex or hippocampal formation lesions on reward assessment in nonhuman primates. Eur J Neurosci. 25:2885–2904.

Maffei A, Haley M, Fontanini A. 2012. Neural processing of gustatory information in insular circuits. Curr Opin Neurobiol. 22:709–716.

Murray EA, Rudebeck PH. 2013. The drive to strive: goal generation based on current needs. Front Neurosci. 7:112.

Ostlund SB, Balleine BW. 2005. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. J Neurosci. 25:7763–7770.

Ostlund SB, Balleine BW. 2007a. The contribution of orbitofrontal cortex to action selection. Ann N Y Acad Sci. 1121:174–192.

Ostlund SB, Balleine BW. 2007b. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. J Neurosci. 27:4819–4825.

Parkes SL, Balleine BW. 2013. Incentive memory: evidence the basolateral amygdala encodes and the insular cortex retrieves outcome values to guide choice between goal-directed actions. J Neurosci. 33:8753–8763.

Parkes SL, Bradfield LA, Balleine BW. 2015. Interaction of insular cortex and ventral striatum mediates the effect of incentive memory on choice between goal-directed actions. J Neurosci. 35:6464–6471.

Parkes SL, Ferreira G, Coutureau E. 2016a. Acquisition of specific response-outcome associations requires NMDA receptor activation in the basolateral amygdala but not in the insular cortex. Neurobiol Learn Mem. 128:40–45.

Parkes SL, Marchand AR, Ferreira G, Coutureau E. 2016b. A time course analysis of satiety-induced instrumental outcome devaluation. Learn Behav. 44:347–355.

Paxinos G, Watson C. 2014. Paxinos and Watson's *The Rat Brain in Stereotaxic Coordinates*. 7th ed. San Diego: Elsevier Academic Press.

Pickens CL, Saddoris MP, Gallagher M, Holland PC. 2005. Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. Behav Neurosci. 119:317–322.

Pickens CL, Saddoris MP, Setlow B, Gallagher M, Holland PC, Schoenbaum G. 2003. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. J Neurosci. 23:11078–11084.

Preuss TM. 1995. Do rats have prefrontal cortex? The rose-woolsey-akert program reconsidered. J Cogn Neurosci. 7:1–24.

Ramirez-Lugo L, Penas-Rincon A, Angeles-Duran S, Sotres-Bayon F. 2016. Choice behavior guided by learned, but not

innate, taste aversion recruits the orbitofrontal cortex. J Neurosci. 36:10574–10583.
Rangel A, Camerer C, Montague PR. 2008. A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci. 9:545–556.
Rhodes SE, 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.
Roberts AC. 2006. Primate orbitofrontal cortex and adaptive behaviour. Trends Cogn Sci. 10:83–90.
Rogan SC, Roth BL. 2011. Remote control of neuronal signaling. Pharmacol Rev. 63:291–315.
Rolls BJ. 1986. Sensory-specific satiety. Nutr Rev. 44:93–101.
Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF. 2008. Frontal cortex subregions play distinct roles in choices between actions and stimuli. J Neurosci. 28:13775–13785.
Rudebeck PH, Mitz AR, Chacko RV, Murray EA. 2013. Effects of amygdala lesions on reward-value coding in orbital and medial prefrontal cortex. Neuron. 80:1519–1531.
Rudebeck PH, Murray EA. 2008. Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. J Neurosci. 28:8338–8343.
Saddoris MP, Gallagher M, Schoenbaum G. 2005. Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. Neuron. 46:321–331.
Samuelsen CL, Gardner MP, Fontanini A. 2012. Effects of cue-triggered expectation on cortical processing of taste. Neuron. 74:410–422.
Schoenbaum G, Nugent SL, Saddoris MP, Setlow B. 2002. Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. Neuroreport. 13: 885–890.
Schoenbaum G, Roesch M. 2005. Orbitofrontal cortex, associative learning, and expectancies. Neuron. 47:633–636.
Schoenbaum G, Setlow B, Nugent SL, Saddoris MP, Gallagher M. 2003. Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. Learn Mem. 10:129–140.
Sherman SM. 2016. Thalamus plays a central role in ongoing cortical functioning. Nat Neurosci. 19:533–541.
Shi CJ, Cassell MD. 1998. Cortical, thalamic, and amygdaloid connections of the anterior and posterior insular cortices. J Comp Neurol. 399:440–468.
Stalnaker TA, Berg B, Aujla N, Schoenbaum G. 2016. Cholinergic interneurons use orbitofrontal input to track beliefs about current state. J Neurosci. 36:6242–6257.
Tang DW, Fellows LK, Small DM, Dagher A. 2012. Food and drug cues activate similar brain regions: a meta-analysis of functional MRI studies. Physiol Behav. 106:317–324.
Ting JT, Daigle TL, Chen Q, Feng G. 2014. Acute brain slice methods for adult and aging animals: application of targeted patch clamp analysis and optogenetics. Methods Mol Biol. 1183:221–242.
Tran-Tu-Yen DA, Marchand AR, Pape JR, Di Scala G, Coutureau E. 2009. Transient role of the rat prelimbic cortex in goal-directed behaviour. Eur J Neurosci. 30:464–471.
Vertes RP. 2004. Differential projections of the infralimbic and prelimbic cortex in the rat. Synapse. 51:32–58.
Wassum KM, Ostlund SB, Maidment NT, Balleine BW. 2009. Distinct opioid circuits determine the palatability and the desirability of rewarding events. Proc Natl Acad Sci U S A. 106:12512–12517.
West EA, Desjardin JT, Gale K, Malkova L. 2011. Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. J Neurosci. 31:15128–15135.
West EA, Forcelli PA, Murnen AT, Mccue DL, Gale K, Malkova L. 2012. Transient inactivation of basolateral amygdala during selective satiation disrupts reinforcer devaluation in rats. Behav Neurosci. 126:563–574.
Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. 2014. Orbitofrontal cortex as a cognitive map of task space. Neuron. 81:267–279.
Yamamoto T. 1984. Taste responses of cortical neurons. Prog Neurobiol. 23:273–315.
Zeeb FD, Winstanley CA. 2013. Functional disconnection of the orbitofrontal cortex and basolateral amygdala impairs acquisition of a rat gambling task and disrupts animals' ability to alter decision-making behavior after reinforcer devaluation. J Neurosci. 33:6434–6443.
Zimmermann KS, Yamin JA, Rainnie DG, Ressler KJ, Gourley SL. 2017. Connections of the mouse orbitofrontal cortex and regulation of goal-directed action selection by brain-derived neurotrophic factor. Biol Psychiatry. 81:366–377.