Article

Neuron

Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation

Highlights

- OFC-DS circuit mediates goal-directed control of actions
- Inhibition of OFC-DS transmission is necessary for habitual action control
- Selective deletion of CB1 receptors in OFC-DS neurons prevents habit formation
- Endocannabinoid-mediated weakening of goal-directed circuits allows for habit formation

Authors

Christina M. Gremel, Jessica H. Chancey, Brady K. Atwood, ..., Karl Deisseroth, David M. Lovinger, Rui M. Costa

Correspondence

lovindav@mail.nih.gov (D.M.L.), rui.costa@neuro.fchampalimaud.org (R.M.C.)

In Brief

Gremel et al. show an endogenous molecular mechanism in an identified corticostriatal circuit that controls shifts in behavioral control between goal-directed and habitual strategies. Habitual action control necessitates the endocannabinoid-mediated weakening of an orbitostriatal circuit that sustains goal-directed control.

Gremel et al., 2016, Neuron *90*, 1312–1324
June 15, 2016
http://dx.doi.org/10.1016/j.neuron.2016.04.043

CellPress

CellPress

Neuron
Article

Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation

Christina M. Gremel,[1,2] Jessica H. Chancey,[1] Brady K. Atwood,[1] Guoxiang Luo,[1] Rachael Neve,[3] Charu Ramakrishnan,[4] Karl Deisseroth,[4] David M. Lovinger,[1,6,*] and Rui M. Costa[5,6,*]

[1]Laboratory for Integrative Neuroscience, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD 20892, USA

[2]Department of Psychology, Neuroscience Graduate Program, University of California San Diego, La Jolla, CA 92093, USA

[3]Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[4]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA

[5]Champalimaud Neuroscience Programme, Champalimaud Institute for the Unknown, Lisbon 1400-038, Portugal

[6]Co-senior author

*Correspondence: lovindav@mail.nih.gov (D.M.L.), rui.costa@neuro.fchampalimaud.org (R.M.C.)

http://dx.doi.org/10.1016/j.neuron.2016.04.043

SUMMARY

Everyday function demands efficient and flexible decision-making that allows for habitual and goal-directed action control. An inability to shift has been implicated in disorders with impaired decision-making, including obsessive-compulsive disorder and addiction. Despite this, our understanding of the specific molecular mechanisms and circuitry involved in shifting action control remains limited. Here we identify an endogenous molecular mechanism in a specific cortical-striatal pathway that mediates the transition between goal-directed and habitual action strategies. Deletion of cannabinoid type 1 (CB1) receptors from cortical projections originating in the orbital frontal cortex (OFC) prevents mice from shifting from goal-directed to habitual instrumental lever pressing. Activity of OFC neurons projecting to dorsal striatum (OFC-DS) and, specifically, activity of OFC-DS terminals is necessary for goal-directed action control. Lastly, CB1 deletion from OFC-DS neurons prevents the shift from goal-directed to habitual action control. These data suggest that the emergence of habits depends on endocannabinoid-mediated attenuation of a competing circuit controlling goal-directed behaviors.

Decision-making requires a balance between flexible and efficient action selection. For normal function, we need to be able to retrieve routine actions quickly and efficiently, and habits serve this purpose. However, the transition between habitual and goal-directed control provides the capacity to perform the same action based on updated consequences. Difficulties with stopping habits and shifting to goal-directed control underlie numerous neuropsychiatric disorders that display impaired decision-making (Dias-Ferreira et al., 2009; Graybiel, 2008; Griffiths et al., 2014), including obsessive-compulsive disorder (OCD) (Burguière et al., 2015; Gillan et al., 2011) and addiction (Belin et al., 2013; Goldstein and Volkow, 2011). However, our understanding of specific molecular mechanisms and circuitry involved in controlling action shifting remains limited.

In everyday decision-making, the contiguity and contingency of actions relative to outcomes can be somewhat predictable. Hence, it is likely that both goal-directed and habitual action strategies are contributing to action control, although to varying degrees. This raises the possibility that decision-making strategies compete for downstream behavioral control. This has been observed in supporting neural circuits, with neural coding of habitual actions in the dorsal medial striatum (DMS), a region known to participate in support of goal-directed control, and neural coding of goal-directed actions in dorsal lateral striatum (DLS), a region known to support habitual processes (Graybiel, 2008; Gremel and Costa, 2013; Hilario et al., 2012; Stalnaker et al., 2010; Thorn et al., 2010; Yin et al., 2004, 2005a, 2006). Thus, disorders such as OCD and addiction may induce a pathology that results in an over-reliance on habitual circuitry in situations in which greater goal-directed control would be more advantageous. Such reliance on habitual action circuitry has been observed following extended and repetitive action training (Smith and Graybiel, 2013; Thorn et al., 2010) and following periods of drug self-administration (Belin and Everitt, 2008; Corbit et al., 2012; Dickinson et al., 2002), with notable exceptions (Colwill and Rescorla, 1985; Samson et al., 2004). These findings underscore the importance of understanding the mechanisms and circuits involved in shifts between these differing action-control strategies.

We have found that the orbital frontal cortex (OFC) is necessary for animals to modulate action value and shift toward goal-directed actions (Gremel and Costa, 2013). We previously developed a within-subject, action-shifting, instrumental lever-press task that allowed us to examine the activity of the same neurons in involved circuits as the animal performed the same action in a goal-directed versus habitual manner. While other findings hypothesized a role for OFC in stimulus-outcome learning (Ostlund and Balleine, 2007; Rolls et al., 1996; Stalnaker et al., 2015), our previous study (Gremel and Costa, 2013) and additional findings (Bradfield et al., 2015; Gourley et al., 2013; Rhodes and Murray, 2013; Stalnaker

et al., 2015) have identified a role for lateral OFC (Gourley et al., 2013; Gremel and Costa, 2013) and medial OFC (Bradfield et al., 2015) in control over goal-directed decision-making that is independent of contextual control yet dependent on the expected action value. Together, along with medial prefrontal cortex (Killcross and Coutureau, 2003), these findings strongly support OFC involvement in goal-directed decision-making. Our previous findings suggested that to shift from goal-directed to habitual actions necessitates a decrease in excitatory transmission from lateral OFC projection neurons, with optogenetic activation selectively increasing the frequency of goal-directed, not habitual, actions and chemogenetic inhibition of lateral OFC projection neurons preventing goal-directed control. Shifts in action control could therefore emerge from changes in the relative weight of OFC and downstream circuits, like DMS, that support goal-directed actions versus competing circuits, like DLS, that support habits (Graybiel, 2008; Gremel and Costa, 2013; Yin et al., 2004, 2005b, 2006). This led us to investigate potential endogenous mechanisms responsible for decreasing OFC transmission, thereby reducing its impact on downstream brain areas and allowing for habitual actions to emerge.

One candidate is the endocannabinoid system. Endocannabinoid signaling at cannabinoid type 1 (CB1) produces synaptic plasticity in corticostriatal circuits that support action strategies (Gerdeman et al., 2002). CB1 receptors are G protein-coupled receptors that activate G proteins of the i of o subclass ($G_{i/o}$) and are expressed in excitatory and inhibitory projection neurons, as well as inhibitory interneurons within corticobasal ganglia circuits controlling actions (Lovinger, 2010). CB1 receptors are largely concentrated on both excitatory and inhibitory presynaptic terminals where, upon activation, they reduce the probability of neurotransmitter release (Kano et al., 2009). Previous work found that activation of the CB1 receptor by endocannabinoids is necessary for habitual learning (Hilário et al., 2007) and that chronic CB1 receptor activation biases toward the use of habitual action strategies (Nazzaro et al., 2012). Therefore, we hypothesized that CB1 receptor activation on OFC projection neurons could be one of the mechanisms gating the shift from goal-directed to habitual action control.

Through the use of Cre recombinase-enabled, cell-specific, and circuit-specific viral approaches in transgenic mice (Fenno et al., 2014), we identify an endogenous molecular mechanism, in a specific cortical-striatal pathway, mediating the transition between goal-directed and habitual action strategies. By selectively deleting CB1 receptors in OFC projection neurons, we identify a role for endocannabinoid modulation of excitatory cortical output in the transition from goal-directed to habitual control. We then show that activity of orbital frontal cortex neurons projecting to dorsal striatum (OFC-DS), and more specifically, transmission at OFC-DS terminals in striatum, is necessary to maintain goal-directed control. Finally, we show that selective deletion of CB1 receptors in OFC-DS neurons is sufficient to block the emergence of habitual behavior. Together, our findings suggest an endogenous mechanism of action shifting, with endocannabinoid-mediated inhibition of OFC-DS circuits allowing for the emergence of habitual control.

RESULTS

To examine cell-type and circuit-specific control over action selection, it is advantageous to examine goal-directed and habitual actions concurrently in the same subject. We therefore used a within-subject, self-paced, instrumental lever-press task that we recently developed (Gremel and Costa, 2013), in which mice readily perform the same action on a similar manipulandum for the same reward using a goal-directed versus a habitual action strategy (Figure 1). In this paradigm, we pair schedules of reinforcement historically used to favor goal-directed or habitual control, random ratio (RR) and random interval (RI), respectively (Adams, 1982; Adams and Dickinson, 1981; Colwill and Rescorla, 1985; Dickinson, 1985), with a particular context. Each day, mice are concurrently trained to press a similar lever in the same location for the same reward (pellets or a 20% sucrose solution) under RI and RR schedules. The other outcome is provided as a control in the home cage (i.e., is not contingent on lever-press behavior). During training, these schedules produce largely similar lever-pressing rates (ANOVA; main effect of day: $F_{8, 192} = 36.40$, $p < 0.001$) (Figure 1B).

To probe the extent to which actions are controlled through goal-directed processes, we conduct a sensory-specific satiety outcome devaluation procedure across 2 days, termed valued day and devalued day (Figure 1C). The valued day provides a control for general effects that satiation may have on lever-press behavior, when mice can prefeed on the home-cage outcome that was not associated with a lever press. Outcome devaluation occurs on the devalued day, when mice can prefeed on the outcome previously earned by lever pressing, thereby decreasing the motivation for that particular outcome. Each day, following prefeeding, mice are placed in each training context and non-rewarded lever pressing is measured. Goal-directed control is sensitive to changes in motivation for the immediate expected outcome, with habitual control less so. Hence, the degree of goal-directedness is assessed through the sensitivity of lever pressing to outcome devaluation (Adams, 1982; Adams and Dickinson, 1981; Colwill and Rescorla, 1985), with a greater reduction in lever pressing in the devalued state compared to the valued state indicative of greater goal-directed action control (and no reduction indicative of habitual control).

Our within-subject task produces goal-directed and habitual action control (Figure 1D). A two-way ANOVA revealed a significant interaction between training context and day ($F_{1, 22} = 10.92$, $p = 0.003$); mice pressed less in the devalued versus the valued state only in the RR context (Bonferroni corrected $p < 0.01$). Mice differentially distributing their lever presses between valued and devalued states in the RR but not the RI training contexts (one-sample t test; RR: $t_{11} = 7.27$, $p < 0.001$) (Figure 1E). There is a shift in the distribution of goal-directedness as measured by a devaluation index [(Valued lever presses − Devalued lever presses)/Total lever presses] (i.e., values closer to 1 are indicative of greater goal-directedness) for each training context, with individual mice showing greater goal-directedness in the RR context than in the RI context (paired t test: $t_{11} = 3.41$, $p = 0.005$) (Figure 1F). Hence, on the same day, mice can readily shift between goal-directed and habitual control over lever pressing for the same outcome.

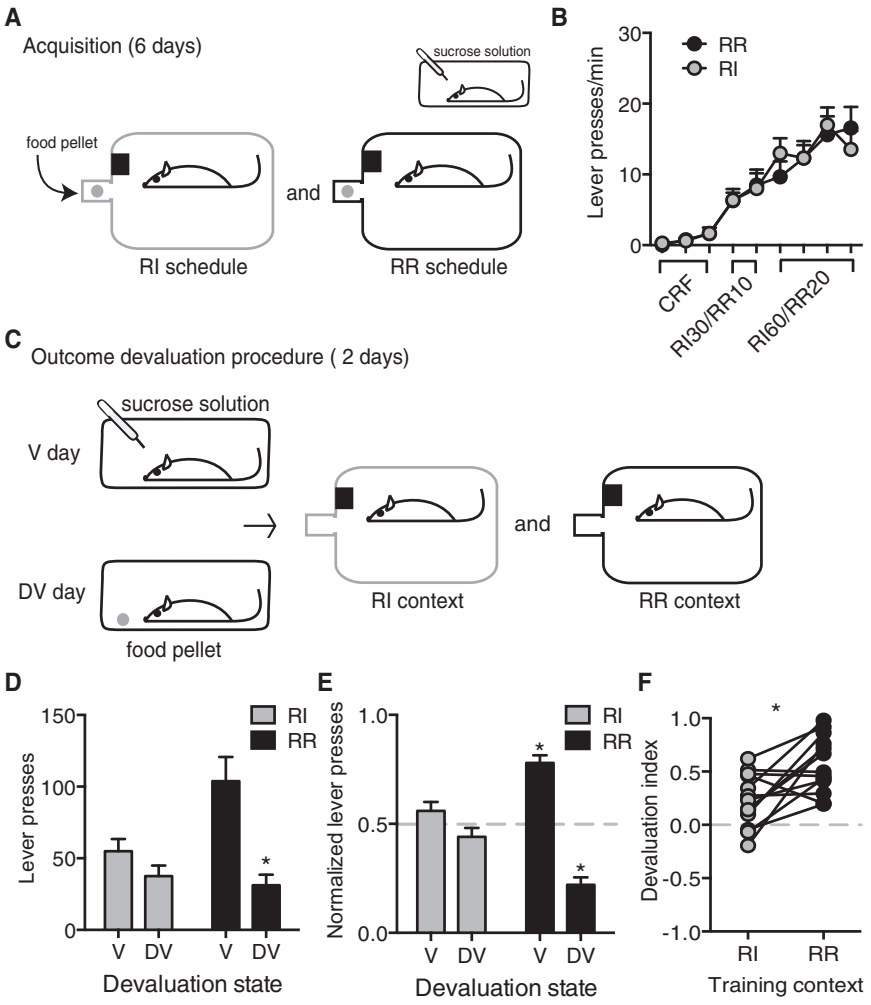

Figure 1. Within-Subject Shifting between Goal-Directed and Habitual Actions

(A) Acquisition schematic of lever pressing for a food outcome under RI and RR schedules of reinforcement. The same mouse is placed in successive order, in two operant chambers distinguished by contextual cues. There, they are trained to press the same lever (e.g., left lever) for the same outcome (e.g., food pellets versus sucrose solution). The bias toward goal-directed actions is generated through use of RR schedules of reinforcement, where the reinforcer is delivered following on average the n lever press (2 days of n = 10, followed by 4 days of n = 20). In contrast, RI reinforcement schedules are used to bias toward use of habitual actions, with the reinforcer delivered following the first lever press after, on average, an interval of t has passed (2 days of t = 30 s, followed by 4 days of t = 60 s). Each day following lever-press training, the other outcome (e.g., sucrose) is provided in the home cage.

(B) Response rate for a control cohort under RI and RR schedules across acquisition.

(C) Schematic of outcome devaluation procedure. On the valued (V) day, mice are fed (1 hr) a control outcome (e.g., sucrose) that they have experienced in their home cage. On the devalued (DV) day, mice are prefed the outcome associated with the lever press (e.g., food pellet). Following prefeeding, mice are placed into the RI and RR contexts, and lever presses are measured for 5 min in the absence of reinforcer delivery.

(D) Lever pressing in V and DV states in RI (gray) and RR (black) contexts.

(E) Distribution of lever presses between V and DV days in RI and RR training contexts.

(F) Within-subject devaluation indexes in previously trained RI and RR contexts, reflecting potential shifts in the magnitude of devaluation.

Individual results and mean ± SEM are shown. *$p < 0.05$.

Deletion of OFC CB1 Receptors Prevents Habitual Control

We first examined the ability to shift between goal-directed and habitual action control following deletion of CB1 receptors in the OFC. We used a viral-vector approach to selectively delete CB1 receptors from OFC neurons in adult mice. *LoxP*-flanked CB1 transgenic mice ($CB1^{flox}$) (Marsicano et al., 2003) and their wild-type littermates were given stereotaxically guided intracranial injections targeted to the lateral OFC of either *AAV9-Ef1a-Cre-eGFP* for widespread Cre recombinase or *AAV9-CamKIICre-eGFP* to specifically target OFC projection neurons (Figure 2A; Figure S1). To confirm the presence or absence of CB1 receptors in terminals of OFC projection neurons after these manipulations (>4 weeks postsurgery), we included Cre-dependent channelrhodopsin (*AAV5-Ef1a-DIO-ChR2-eYFP*) in some of the OFC injections and examined the ability of the CB1 agonist WIN55,212 to reduce light-evoked excitatory transmission from OFC terminals onto medium spiny neurons (MSNs) in the DS via whole-cell patch clamp electrophysiology.

Consistent with previous findings showing extensive synaptic connectivity between OFC and striatum in mice (Figure 2B) (Pan et al., 2010; Wall et al., 2013), we detected optically evoked excitatory postsynaptic currents (oEPSCs) in MSNs in DS (Figure 2C). As expected, application of WIN55,212 reduced the amplitude of oEPSPs in control mice (unpaired t test: $t_{14} = 3.55$, $p = 0.003$), with a one-sample t test against baseline showing reduced relative EPSCs in control mice ($t_5 = 5.42$, $p = 0.003$) (Figures 2C and 2D; Figure S1E), indicating intact CB1 receptor function in OFC presynaptic terminals. In contrast, WIN55,212 did not reduce oEPSCs in virally injected $CB1^{flox}$ mice ($p > 0.05$), confirming deletion of CB1 receptors from OFC presynaptic terminals.

Adult control mice (wild-type littermates injected with virus), $CB1^{flox}$ mice with CB1 receptors deleted from OFC neurons ($CB1^{flox}$Cre), and $CB1^{flox}$ mice with CB1 receptors deleted from OFC CamKII-expressing neurons ($CB1^{flox}$CamKII Cre) were trained to lever press for the same food reward under both RI and RR schedules of reinforcement (Figure 2E). Deletion of CB1 receptors from OFC neurons had a small effect on acquisition of lever-press behavior; $CB1^{flox}$Cre mice and $CB1^{flox}$CamKII Cre mice made slightly more lever presses as the RR requirement increased (Figure S2). $CB1^{flox}$Cre mice and $CB1^{flox}$CamKII

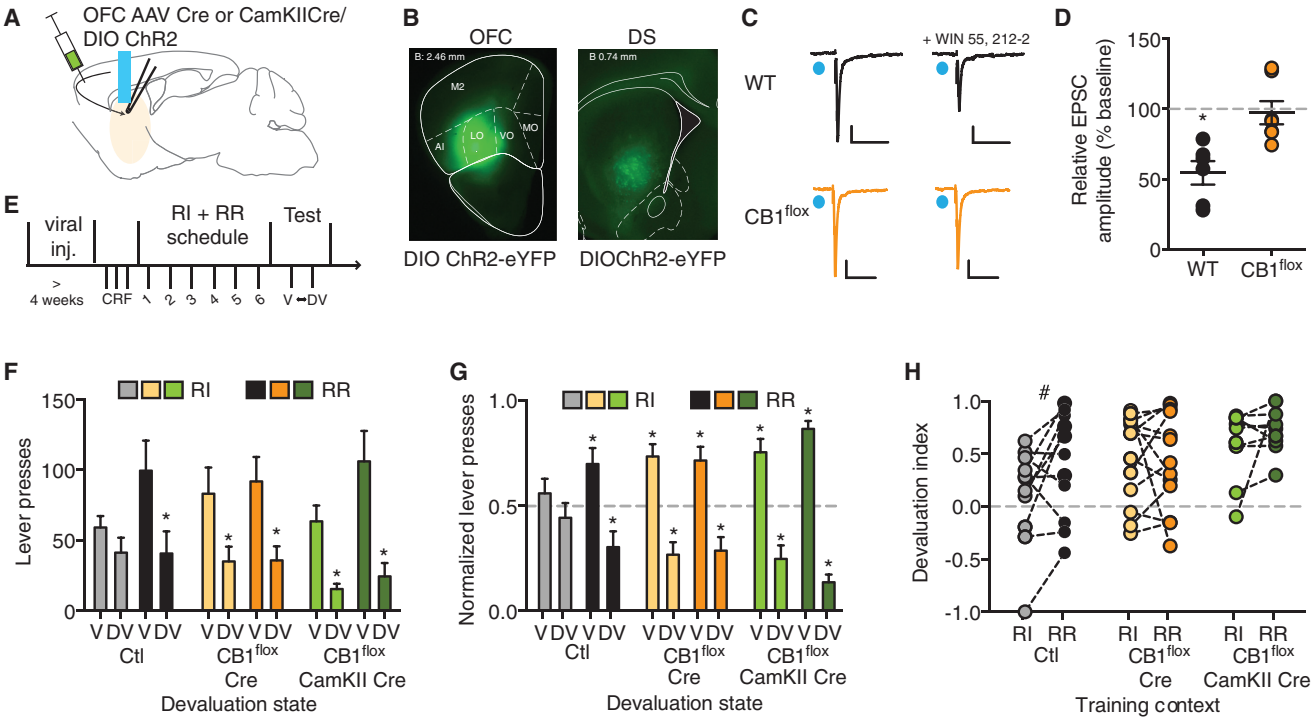

Figure 2. Deletion of CB1 Receptor from OFC Neurons Impairs Habitual Action Control

(A) Schematic of viral strategy and ex vivo physiological assessment.

(B) Cre-dependent ChR2 eYFP detected at OFC injection site (left) and downstream DS (right) in CB1^{flox}CamKII Cre mice.

(C) Representative traces showing assessment of oEPSCs in DS MSNs, with the blue circle indicating the light pulse (473 nm wavelength, 5 ms) in CB1^{flox}CamKII Cre mice and wild-type littermates. WIN55,212, is a CB1 receptor agonist. Scale bars: y axis = 50 pA, x axis = 100 ms.

(D) Relative amplitude of DS MSN oEPSCs following WIN55,212 application (10 min, 1 μM) differed between wild-type and CB1^{flox}Cre or CB1^{flox}CamKII Cre mice.

(E) Experimental design schematic.

(F) Lever presses made following outcome devaluation procedures in valued (V) and devalued (DV) states across RI and RR training contexts for control (Ctl) mice, CB1^{flox}Cre mice, and CB1^{flox}CamKII Cre mice.

(G) Normalized lever presses during outcome devaluation testing, showing the distribution of lever presses between V and DV states in the different training contexts (RI and RR) across the different treatment groups.

(H) Shifts in outcome devaluation index between RI and RR training contexts for the different treatment groups.

Individual results and mean ± SEM are shown. *p < 0.05, #p = 0.08. See also Figures S1 and S2.

Cre mice were able to distinguish between RI and RR schedules, as evidenced by maintained sensitivity to differential feedback functions produced by each schedule (Figures S2G–S2I).

To evaluate the degree to which mice used the expected outcome value to control decision-making, we performed an outcome devaluation test. Control mice selectively reduced lever pressing following outcome devaluation in the RR context but not the RI context (Figure 2F). Repeated-measures ANOVA showed a Context × Day interaction in control mice and reduced responding on the devalued day only in the RR context ($F_{1, 20}$ = 2.79, p = 0.05; Bonferroni corrected p < 0.05). In contrast, CB1^{flox}Cre mice and CB1^{flox}CamKII Cre mice made fewer lever presses in the devalued compared to the valued state in previously trained RI and RR contexts (Figure 2F). A main effect of day was observed in CB1^{flox}Cre mice ($F_{1, 24}$ = 18.08, p = 0.0003) and CB1^{flox}CamKII Cre mice ($F_{1, 16}$ = 41,47, p < 0.0001), with reduced lever pressing in devalued versus valued states in both RI and RR contexts (Bonferroni corrected p values (ps′) < 0.05). Furthermore, a one-sample t test (against chance or 0.5) of normalized lever presses between valued and devalued states in each training context showed that control mice were only sensitive to devaluation in the RR context (RI: t_{10} = 0.82, p = 0.43; RR: t_{10} = 2.67, p = 0.025). In contrast, CB1^{flox}Cre mice and CB1^{flox}CamKII Cre mice were sensitive to devaluation in previously trained RI and RR contexts (CB1^{flox}Cre mice, RI: t_{12} = 3.96, p = 0.002, and RR: t_{12} = 3.33, p = 0.006; CB1^{flox}CamKII Cre mice, RI: t_8 = 3.96, p = 0.003, and RR: t_8 = 9.83, p < 0.0001) (Figure 2G).

Finally, to evaluate whether CB1 deletion altered the within-subject shift in goal-directedness, we performed a paired t test on the devaluation index for each group to compare the degree to which each mouse was goal-directed in the RI versus the RR context. Although all mice consumed similar amounts of food during sensory-specific satiation procedures (Figure S2F), control mice tended to show greater sensitivity to outcome devaluation in the RR than RI context (t_{20} = 2.36, p = 0.08) (Figure 2H). In contrast, CB1^{flox}Cre mice and CB1^{flox}CamKII Cre mice showed similar sensitivity to outcome devaluation in both RI and RR contexts (CB1^{flox}Cre mice: t_{24} = 0.21, p = 0.83; CB1^{flox}CamKII Cre mice: t_{16} = 1.07, p = 0.30). These findings are in accordance

with additional experiments using an inducible forebrain-specific Cre-recombinase approach, which suggests a role for OFC CB1 receptors in shifting toward habitual action control (Figure S3). Together, these data strongly suggest that deletion of CB1 receptors from OFC neurons, and specifically, CamKII-expressing neurons in adulthood, prevents a shift to habitual action control and renders mice reliant on goal-directed processes in both contexts.

Attenuation of OFC-DS Projection Activity Results in Habitual Control

We previously showed that efficacy of excitatory transmission from OFC CamKII-expressing neurons to their targets maintains goal-directed control, with chemogenetic inhibition of OFC projection neurons leaving mice reliant on habitual decision-making processes (Gremel and Costa, 2013). However, OFC neurons send projections widely across the cortex, striatum, and midbrain areas (Hoover and Vertes, 2011; Pan et al., 2010; Schilman et al., 2008). Given dorsal-striatal involvement in goal-directed and habitual actions, we first tested whether the activity of OFC-DS neurons controls the shift from goal-directed to habitual actions. We took a combined retrograde and chemogenetic viral approach to silence activity of OFC-DS neurons during outcome devaluation (Armbruster et al., 2007; Fenno et al., 2014; Stachniak et al., 2014; Znamenskiy and Zador, 2013). We selectively expressed the inhibitory G_i-coupled human muscarinic 4 designer (hM4D) receptor in OFC-DS neurons and performed outcome devaluation procedures in the presence of the selective exogenous ligand clozapine-N-oxide (CNO), thereby inhibiting OFC-DS neurons.

Adult wild-type mice were injected with the retrograde herpes simplex virus 1 carrying Cre recombinase (*hEF1α-eYFP-IRES-cre*; HSV-1 Cre) or enhanced yellow fluorescent protein (eYFP) (*hEF1α -EYFP-IRES*; HSV-1 YFP) stereotaxically targeted to the DMS, where we had previously observed OFC fiber tracts and recorded oEPSCs in MSNs (Figure 3A; Figure S4). In the same surgery, mice were given an injection of *AAV5-Ef1a-DIO-hM4D-mCherry* into lateral OFC (Figure 3A). This combination of viruses resulted in expression of hM4D only in OFC-DS (Figure 3C). We did not observe hM4D-induced mCherry fluorescence following *hEF1α -EYFP-IRES* control injections into striatum (control mice) (Figure S4). We have previously shown that CNO activation of hM4D receptors on OFC projection neurons leads to a reduction in OFC activity in vivo (Gremel and Costa, 2013). To verify that CNO activation of hM4D receptors inhibited activity of OFC-DS neurons (Stachniak et al., 2014), we performed whole-cell patch recordings of visually identified hM4D-expressing OFC-DS neurons. We observed hyperpolarization of OFC neurons with bath application of CNO (Figure S4D); CNO application blocked current-evoked action potential firing, whereas vehicle application did not (Figures 3D and 3E).

We trained control and OFC-DS hM4D mice on our within-subject lever-press task. Following acquisition, we performed the outcome devaluation procedure in the presence of CNO (1 mg/kg intraperitoneally) (Figure 3B). Again, control mice showed reduced responding in the devalued state in the RR context but not the RI context (repeated-measures ANOVA, Context × Day: $F_{1,12} = 0.8$, $p = 0.37$; preplanned paired t test, RR: $t_{12} = 2.44$, $p = 0.05$) (Figure 3F). In contrast, OFC-DS hM4D mice did not reduce responding in the devalued state in either the RI or the RR context ($F_{1, 18} = 0.98$, $p = 0.34$; $ps' > 0.05$). Furthermore, we found that while control mice differentially distributed lever presses between valued and devalued states only in the RR context (RI: $t_6 = 1.69$, $p = 0.14$; RR: $t_6 = 5.00$, $p = 0.002$) (Figure 3G), OFC-DS hM4D mice showed similar lever-press distributions in each context (RI: $t_9 = 0.11$, $p = 0.91$; RR: $t_9 = 1.28$, $p = 0.23$). We then examined whether hM4D activation and silencing of OFC-DS altered the shift in devaluation index between RI and RR contexts. We found that while control mice were overall more goal-directed in the RR than the RI training context, this was not the case for OFC-DS hM4D mice (Figure 3H). A paired t test of devaluation indexes in the RI versus the RR context revealed a significant shift in control mice ($t_{12} = 2.33$, $p = 0.03$) that was not present in OFC-DS hM4D mice ($t_{18} = 0.44$, $p = 0.66$). The observed increase in the overall level of lever pressing in OFC-DS hM4D mice (Figure 3F) raises the possibility that increased locomotor activity interfered with the ability to show outcome devaluation. Correlational analyses did not support this hypothesis (Figures S4E and S4F), nor did CNO induce a general increase in locomotor activity in control or OFC-DS hM4D mice (Figure S4G).

However, because deep-layer cortical projection neurons have collateral projections (Shepherd, 2013), systemic administration of CNO could result in attenuated collateral activity of OFC-DS neurons in other downstream terminal areas. To directly examine the contribution of OFC terminals in the DS, we infused CNO directly into the more medial DS before outcome devaluation procedures. We implanted cannulae targeted to the DMS in adult wild-type mice that had been previously injected with *AAV5-Ef1a-Cre-eGFP* into DS (Rothermel et al., 2013) and *AAV5-Ef1a-DIO-hM4D-mCherry* into lateral OFC (Figure 4C; Figure S4). Following recovery and subsequent lever-press training on RI and RR schedules, we microinjected CNO (300 μM; n = 6) or saline (n = 6) into the DS before devaluation testing.

We found that the handling procedures associated with microinjections before devaluation testing were sufficient to disrupt habitual control in saline-infused mice, with saline mice showing strong goal-directed control over actions in both training contexts (ANOVA; no interaction, main effect of day: $F1, 10 = 29.30$, $p < 0.001$) (Figure 4D). Despite this disruption, intra-DS infusions of CNO resulted in strong habitual control in both RI and RR training contexts (no interaction, no main effects: $ps' > 0.05$). Control mice differentially distributed their lever presses between valued and devalued states in both RI and RR contexts (RI: $t_5 = 22.82$, $p < 0.001$; RR: $t_5 = 15.18$, $p < 0.001$), whereas CNO-treated mice did not (RI: $t_5 = 1.46$, $p > 0.05$; RR: $t_5 = 1.07$, $p > 0.05$) (Figure 4E). Neither group showed a shift in devaluation index ($ps' > 0.05$) (Figure 4F). A planned comparison performed on devaluation indexes between groups revealed greater goal-directedness in the saline- versus CNO-treated mice (main effect of treatment group: $F1, 10 = 26.58$, $p < 0.01$). Hence, saline mice were goal-directed, while CNO-induced attenuation of OFC-DS transmission left mice reliant on habitual strategies to control lever pressing.

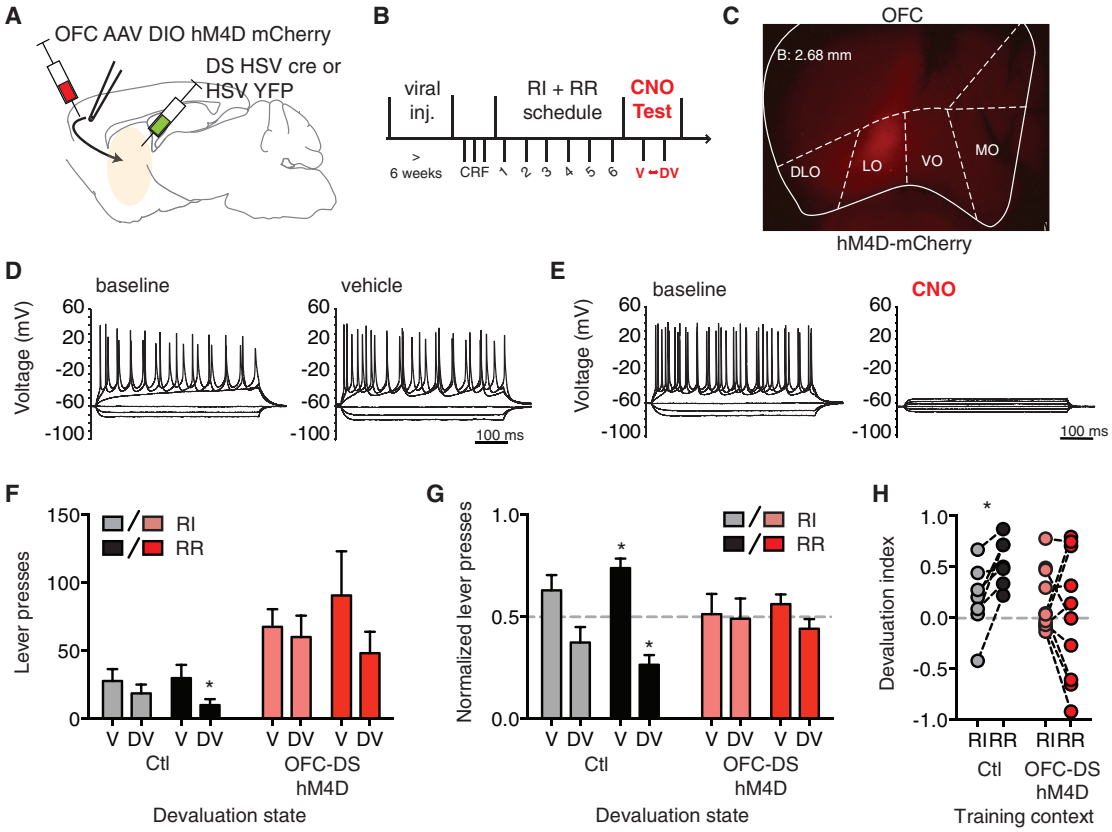

Figure 3. Activity in OFC-DS Neurons Is Necessary for the Shift to Goal-Directed Action Control

(A) Schematic of combinatorial retrograde and AAV viral strategy.

(B) Schematic of experimental design with devaluation testing performed following CNO administration.

(C) hM4D mCherry expression in OFC.

(D and E) Representative traces showing the ability of injected current (−200 to +300 pA, 100 pA steps) to evoke an action potential in OFC neurons (baseline) and following vehicle (DMSO; D) and CNO (10 μM; E) application.

(F) Lever presses during outcome devaluation testing across valued (V) and devalued (DV) states in RI and RR training contexts.

(G) Normalized lever presses during outcome devaluation testing showing the distribution of lever presses between V and DV states in the previously trained RI and RR contexts.

(H) Devaluation index for each group of mice in the previously trained RI and RR contexts.

Individual results and mean ± SEM are shown. *p < 0.05. See also Figure S3.

The lack of goal-directed control following CNO treatment is not explained by disrupted response inhibition, because in a separate test, systemic CNO treatment did not prevent satiety-induced responding reduction in an open-ended fixed ratio schedule (Figures S4J–S4L). Instead, our findings are in line with reports of alterations in OFC-DS activity in repetitive action pathologies, such as OCD (Burguière et al., 2013) and alterations that may underlie previous reports relating habitual responding and OCD (Gillan et al., 2011). Together, they suggest that hM4D activation and subsequent attenuation of OFC-DS neuron transmission selectively disrupt the ability for goal-directed strategies to control lever-press responding.

CB1 Receptor-Mediated Attenuation of OFC-DS Activity Is Critical for Habitual Control

The previous data suggested that activity of OFC-DS neurons is critical for goal-directed behavior. We therefore reasoned that attenuation of transmission in this pathway via CB1 receptor activation could be critical for habit formation. We next investigated whether endogenous CB1 receptor activation and subsequent reduction in glutamate release from OFC-DS neurons was necessary for the shift from goal-directed to habitual action control. We used a site-specific targeted combinatorial viral approach (Fenno et al., 2014) in transgenic mice to selectively delete CB1 receptors in only OFC-DS. $CB1^{flox}$ mice and their wild-type littermates were injected with a retrograde virus carrying flipase *hEF1α-eYFP-IRES-flp* (HSV-1 fp) targeted to the same DMS region of observed OFC-DS connections (Figure 5A; Figure S5). In the same surgery, we stereotaxically injected *AAV8-Ef1a-FD-mCherry-p2A-Cre* targeted to the lateral OFC, with Cre recombinase dependent on the presence of flipase (adeno-associated virus [AAV] fp-Cre). The aim of this site-specific viral targeting was to limit Cre-recombinase expression to OFC-DS neurons and CB1 deletion in OFC-DS

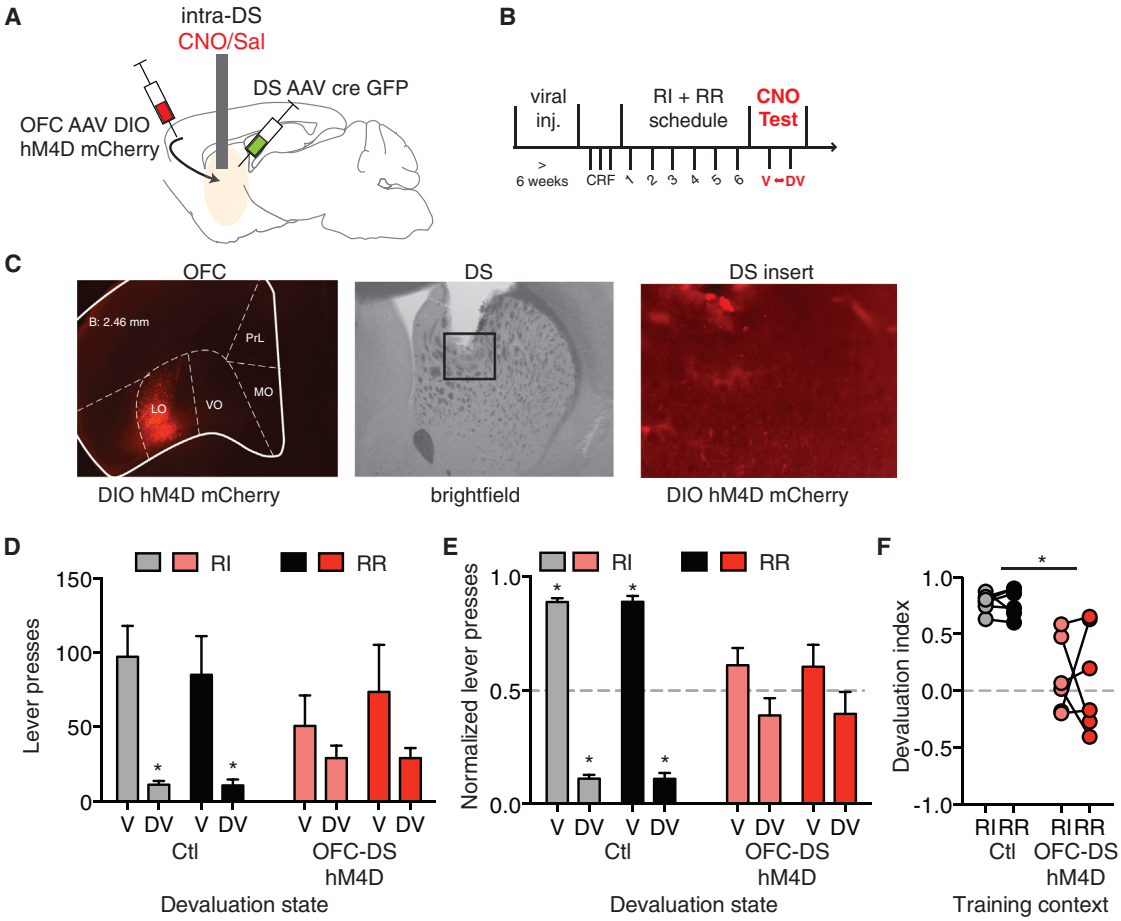

Figure 4. Attenuating OFC-DS Transmission Disrupts Goal-Directed Control

(A) Schematic of combinatorial AAV viral strategy.

(B) Schematic of experimental design with devaluation testing performed following intracranial CNO or saline administration.

(C) DIO hM4D mCherry expression in OFC (left), cannula placement within DS (center), and DS insert from block outline in center panel showing DIO hM4D mCherry fiber expression in DS (right).

(D) Lever presses during outcome devaluation testing across valued (V) and devalued (DV) states in RI and RR training contexts.

(E) Normalized lever presses during outcome devaluation testing showing the distribution of lever presses between V and DV states in the previously trained RI and RR contexts.

(F) Devaluation index for each group of mice in the previously trained RI and RR contexts.

Individual results and mean ± SEM are shown. *$p < 0.05$. See also Figure S4.

neurons of CB1^{flox} mice (OFC-DS CB1^{flox}) but not in wild-type littermates (control). We included *AAV5-Ef1a-DIO-ChR2-eYFP* (AAV DIO ChR2) to functionally assess CB1 receptor deletion (Figures 5A and 5B). We were able to reliably measure light-evoked changes in membrane potential of OFC somata (Figure 5C), indicating successful Cre-recombinase activity. Although more challenging due to the relative sparseness of terminals, we were able to perform whole-cell patch recordings and evoke oEPSCs in a subset of MSNs from wild-type control and CB1^{flox} mice (Figure 5E). WIN55,212 reduced the amplitude of oEPSCs in wild-type but not CB1^{flox} mice (Figures 5E and 5F). These data suggest that our site-specific targeted combinatorial viral approach in transgenic mice was successful at deleting CB1 receptors in terminals of OFC-DS neurons.

Deletion of CB1 receptors from OFC-DS neurons did not alter acquisition of lever-press behavior under RI or RR reinforcement schedules, and mice were able to distinguish between schedules (Figure S5). During outcome devaluation procedures, control mice showed reduced lever pressing in the devalued state in the RR context but not the RI context (repeated-measures ANOVA; Context × Day: $F1, 12 = 5.47$, $p = 0.04$; Bonferroni corrected ps′ < 0.05) (Figure 5G). However, OFC-DS CB1^{flox} mice reduced lever pressing in both RI and RR contexts in the devalued state (no interaction, main effect of day: $F1, 18 = 29.72$, $p < 0.0001$; Bonferroni corrected ps′ < 0.05). Furthermore, OFC-DS CB1^{flox} mice differentially distributed their lever pressing between valued and devalued states in both RI and RR contexts (RI: $t_{12} = 6.14$, $p < 0.0001$; RR: $t_{12} = 2.45$, $p < 0.05$) (Figure 5H). Controls only differentially distributed pressing in the RR context

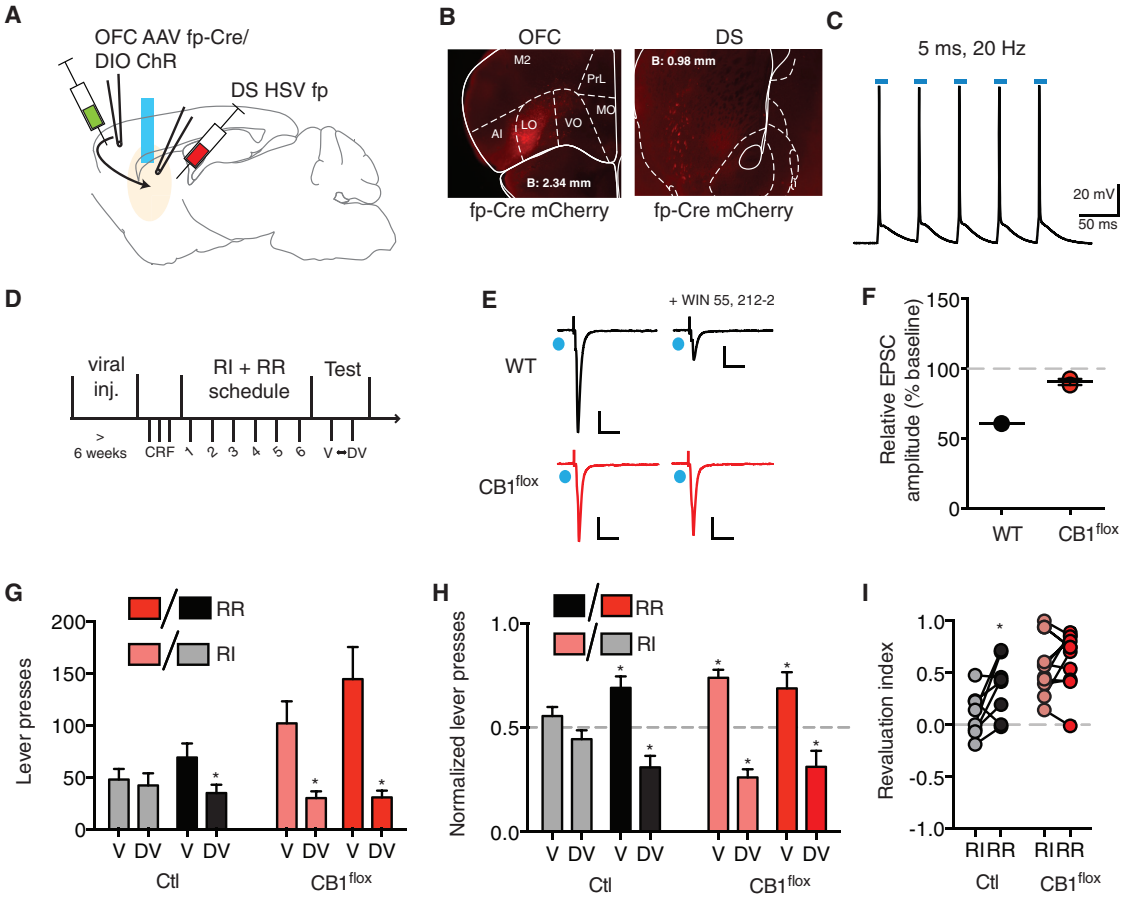

Figure 5. Deletion of CB1 Receptors in OFC-DS Neurons Prevents Habitual Control over Actions

(A) Schematic of combinatorial viral strategy and ex vivo physiological assessment.

(B) Flipase (fp)-dependent Cre-mCherry detected in OFC (left) and in downstream DS (right).

(C) Representative traces showing ChR2-mediated firing of an OFC neuron.

(D) Schematic of experimental design.

(E) Representative traces showing assessment of DS MSN oEPSCs in a subset of OFC-DS CB1^{flox} mice (n = 2) and wild-type littermate mice (Ctl; n = 1) in the absence (left) and presence (right) of WIN55,212. Scale bars: y axis = 100 pA, x axis = 20 ms.

(F) Relative amplitude of DS MSN oEPSCs following WIN55,212 application in Ctl and OFC-DS CB1^{flox} mice.

(G) Lever presses in valued (V) and devalued (DV) states in RI and RR training contexts.

Individual results and mean ± SEM are shown.

(H) Normalized lever presses during outcome devaluation testing showing the distribution of lever presses across V and DV states in the different training contexts (RI and RR).

(I) Devaluation index plotted within subject for RI and RR training contexts for control mice and OFC-DS CB1^{flox} mice.

Individual results and mean ± SEM are shown. *p < 0.05. See also Figure S5.

(RI: t_7 = 1.3, p = 0.24; RR: t_7 = 3.45, p = 0.011). Finally, control mice increased their devaluation index between RI and RR contexts (t_{14} = 2.11, p = 0.05) (Figure 5I), indicative of a shift in the degree of goal-directedness; OFC-DS CB1^{flox} mice did not show such a shift (t_{20} = 0.49, p = 0.63). This suggests that habitual action control involves CB1 receptor-mediated inhibition of the OFC-DS circuit that supports goal-directed actions.

DISCUSSION

The data presented in this study uncover a novel mechanism for shifting toward habitual control. Our results show that competition between DMS and DLS circuits supporting goal-directed and habitual action, respectively, is strongly influenced by the gating of incoming transmission from OFC. By targeting molecular mechanisms in a cell-type, circuit-specific, and projection-specific manner, we identify an endogenous mechanism that underlies gating shifts between goal-directed and habitual action control in the same subject.

In the same mouse, we were able to examine the loss of goal-directed action control in one context while goal-directed control was intact in the remaining context. We used differing schedules of reinforcement to bias differential action control. The different RI and RR schedules we used to bias different action-control

strategies do not produce differences in the macrostructure or microstructure of lever pressing or in the rate across a session (i.e., interval-induced scalloping of lever pressing) (Gremel and Costa, 2013). This within-subject approach to investigate habitual action control differs from that of an extending training approach. In some circumstances (Colwill and Rescorla, 1985), extensive experience with the action-outcome relationship through an extending training approach shifts the control of a subject's responding from initial goal-directed control to habitual control (Dickinson, 1985). Hence, in the within-subject procedure used in the current experiments, it is not the animal that is habitual or goal-directed; it is the action control within a particular context that is habitual or goal-directed (Figure 1). Although there may be subtle differences, habitual control biased through use of both extended training or RI schedules recruits the DLS (Barnes et al., 2005; Dias-Ferreira et al., 2009; Gremel and Costa, 2013; Hilario et al., 2012; Smith and Graybiel, 2013; Thorn et al., 2010; Yin et al., 2004, 2006; Yin and Knowlton, 2006), whereas goal-directed control biased through the use of RR depends on the DMS (Barnes et al., 2005; Dias-Ferreira et al., 2009; Gremel and Costa, 2013; Hilario et al., 2012; Smith and Graybiel, 2013; Thorn et al., 2010; Yin et al., 2004, 2005a, 2005b, 2006; Yin and Knowlton, 2006). Using this within-subject procedure, we took a chemogenetic approach to inhibit the activity of OFC-DS neurons and found that inhibition of OFC-DS activity via activation of the $G_{i/o}$-coupled hM4D receptor specifically during probe test sessions left the subject reliant on habit circuitry, whereas goal-directed circuits are normally favored (Figure 3). Furthermore, specific attenuation of transmission at OFC terminals in DS prevented use of goal-directed strategies for action control (Figure 4). This suggests that increasing $G_{i/o}$ modulation at OFC terminals in DS contributes to the emergence of habitual action control. Whether this mechanism is also responsible for the shift to habitual control following extended training remains to be investigated.

Previous work had implicated an endogenous $G_{i/o}$-coupled receptor, the CB1 receptor, in habitual action control (Hilário et al., 2007). Strikingly, CB1 receptor antagonism during acquisition alone prevented use of habitual action control. This suggests that CB1 receptors are important for the learning-induced plasticity that is necessary for habitual control. Given the widespread expression and increasing gradient of CB1 found in medial to lateral striatum (Kano et al., 2009), it seemed likely that CB1 receptor modulation of habitual control occurred through actions in corticostriatal circuits. Most CB1 receptors are found on presynaptic inhibitory synapses within striatum (Mátyás et al., 2006; Uchigashima et al., 2007). The expression on these terminals likely accounts for the higher levels of expression in DLS than in DMS. However, CB1 receptors are also found on cortical projections into striatum, where they are localized to terminals (Uchigashima et al., 2007), and endocannabinoids are released postsynaptically in an activity-dependent manner (Gerdeman et al., 2002; Uchigashima et al., 2007). These endocannabinoids bind CB1 receptors presynaptically to serve as a powerful modulator of glutamate release (Gerdeman and Lovinger, 2001). We targeted these corticostriatal CB1 receptors and found that prior to training, the specific deletion of CB1 receptors in OFC neurons, and in particular OFC-DS neurons, resulted in a loss of habitual action control (Figures 2 and 5). Together, our data suggest that habitual control over an action strategy depends on $G_{i/o}$-coupled receptor-mediated attenuation of a competing circuit controlling goal-directed behaviors.

The existence of parallel associative and sensorimotor circuits that control disparate action strategies suggests that perturbations to those controlling goal-directed strategies may result in an abnormal reliance on habit circuitry. The reported reduction in task-induced activation of OFC-DS circuits (Remijnse et al., 2006) and the greater reliance on habitual processes (Gillan et al., 2011) observed in OCD patients raise the hypothesis that a pathology involving reduced activity of OFC-DS circuits is present with OCD. In a genetic mouse model of OCD, an acute increase in transmission specifically at OFC-DS terminals reduced compulsive grooming behavior (Burguière et al., 2013). In addition, chronic stimulation of OFC inputs into striatum resulted in development of a repetitive grooming pathology in wild-type mice (Ahmari et al., 2013). Furthermore, both clinical and preclinical work have implicated dysfunction of OFC circuits in addiction disorders (Goldstein and Volkow, 2011; Schoenbaum and Shaham, 2008) that involve impaired decision-making. Chronic administration of the main psychoactive ingredient in cannabis, Δ9-THC, also biases toward use of habitual strategies, and chronic Δ9-THC, as well as alcohol exposure, disrupts endocannabinoid modulation in DS (DePoy et al., 2013; Nazzaro et al., 2012). Thus, loss of endocannabinoid-mediated plasticity in OFC-DS circuits could contribute to addiction (Belin et al., 2013; Goldstein and Volkow, 2011) and OCD (Burguière et al., 2015; Gillan et al., 2011). We also observed increased response (lever pressing) rates following attenuation of OFC-DS neuron activity via $G_{i/o}$-coupled hM4D receptor activation and following deletion of CB1 receptors from OFC-DS neurons that were independent of any effects on outcome devaluation. While the increased response rates may reflect a lack of competition between circuits, they cannot contribute to action control; we found that $G_{i/o}$-coupled hM4D receptor activation led to habitual control, while CB1 deletion from OFC-DS neurons led to loss of habitual control. Furthermore, the observed increased response rates are not due to lack of inhibitory control, because both show sensitivity to satiety. Our studies do suggest OFC-DS circuits underlie a critical component of action control that is disturbed in disorders affecting decision-making control over actions.

With an impaired habit-breaking phenotype present in numerous psychiatric disorders (Griffiths et al., 2014), and widespread use of the abused drugs in the general populace, our findings suggest that therapeutic targeting of the endocannabinoid system is a viable option for restoration of goal-directed control.

EXPERIMENTAL PROCEDURES

All experiments were approved by the Animal Care and Use Committees of the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the University of California San Diego, and performed in accordance with NIH guidelines. Gene-targeted mice $Cnr1^{loxP/loxP}$ were obtained (Marsicano et al., 2003) and bred in-house. Male and female mice were housed in groups of one to four, with mouse chow and water ad libitum unless stated otherwise, and were kept on a 12 hr or 14/10 hr light/dark cycle. All surgical, electrophysiological, and behavioral experiments were performed during the light portion of the cycle.

Deletion of CB1 Receptor from OFC Neurons

Adult CB1^{flox} mice and their wild-type littermates were given stereotaxically guided intracranial injections targeted at the OFC (anterior [A], 2.70 mm; medial [M], ±1.75 or 1.65 mm; ventral [V]: −2.25 mm; Bregma) of either *AAV9-Cre-eGFP* or *AAV9-CamKIICre-eGFP* (200 nL per side), along with *AAV5-DIO-ChR2-eYFP* (100 nL per side; University of Pennsylvania vector core). At least 4 weeks following viral injection, mice were trained on the within-subject instrumental task or were used for whole-cell patch clamp physiology. Virus spread in each mouse put through behavioral procedures was examined under a fluorescence microscope (Olympus MVX10); mice with fluorescence outside the OFC inclusion area were excluded from the final data analyses (n = 9). Final group numbers from a total of three replicates are as follows: control, n = 12; CB1^{flox}Cre, n = 13; and CB1^{flox}CamKII Cre, n = 9.

Chemogenetic Inhibition of OFC-DS Neurons

C57BL/6J mice each received two bilateral injections, one targeting the DS and the other the OFC. In one experiment, *hEF1α-eYFP-IRES-cre* or *hEF1α-EYFP-IRES* (Massachusetts Institute of Technology [MIT] vector core) was injected into the DS (300 nL/side; A, 0.5 mm; M, ±1.25 mm; V, −3.25 mm; Bregma), and *AAV5-DIO-hM4D-mCherry* was injected into the OFC (200 nL/side; coordinates same as earlier; Gene Therapy Vector Core at the University of North Carolina). In a second experiment, mice were given an infusion of *AAV5-Cre-GFP* (200 nL; Gene Therapy Vector Core at the University of North Carolina) into the DS and *AAV5-DIO-hM4D-mCherry* (200 nL) into OFC. These mice were also implanted with bilateral cannulae (Plastics One) targeted to the DS (A, 0.5 mm; M, ±1.25 mm; V, −2.75 mm; Bregma). Following at least 4 weeks of recovery, mice were trained in the within-subject task. In experiments with systemic CNO administration, each day 60 min before devaluation testing, mice were given an intraperitoneal injection of CNO (1 mg/kg, 10 mL/kg). The mice with cannulae were lightly anesthetized, dummy cannulae were removed, and injectors targeting DS (−3.25 mm) were lowered. Then, 15–30 min before the onset of the devaluation procedure, mice were give 300 nL intracranial injections of either 0.9% isotonic saline or CNO (300 μM) at a rate of 100 nL/min. Following devaluation testing, these mice also underwent an additional lever-press test. Mice were pretreated (15 min) with systemic saline or CNO (1 mg/kg, 10 mL/kg) before access to the previously earned outcome on a fixed ratio of 8. To confirm successful viral infection and expression of hM4D receptors, a subset of injected mice were euthanized and whole-cell patch clamp recordings were used to verify CNO-induced suppression of OFC-DS neuronal activity. Viral spread was observed under a fluorescence microscope (Olympus MVX10), and mice with no expression or significant expression outside of OFC inclusion boundaries (n = 5) or with misplaced cannulae or infection at the site of intracranial injection (CNO treated, n = 4; saline treated, n = 3) were excluded. The final group numbers from two systemic replicates are as follows: control, n = 7, and hM4D, n = 10. The final group numbers from one intracranial replicate are as follows: saline treated, n = 6, and CNO treated, n = 6.

Deletion of CB1 Receptors from OFC-DS Neurons

To selectively delete CB1 receptors in OFC-DS neurons in adult mice, we bilaterally injected CB1^{flox} and wild-type littermates with *hEF1α-eYFP-IRES-flpo* (300 nL/side; MIT vector core) into DS, and *AAV8-Ef1a-FD-mCherry-p2A-Cre* (200 nL per side; courtesy of Dr. Deisseroth, Stanford University) and *AAV5-DIO-ChR2-eYFP* (100 nL/side; University of Pennsylvania Vector Core) into OFC. At least 6 weeks following viral injection, mice were trained on the within-subject procedure. Histological examination was done to confirm viral expression, and mice with no mCherry/YFP expression were excluded from the final behavioral analyses (n = 9). A subset of mice was used for ex vivo physiological assessment via whole-cell patch clamp. Final behavioral group numbers from two replicates are as follows: control, n = 7, and OFC-DS CB1^{flox}, n = 10.

Lever-Press Training

All behavioral training and testing took place as previously described (Gremel and Costa, 2013). In brief, mice were placed in operant chambers in sound attenuating boxes (Med-Associates) in which they pressed a single lever (left or right) in a self-paced manner for an outcome of either regular chow pellets (20 mg pellet per reinforcer, Bio-Serve formula F05684) or sucrose solution (20–30 μL of 20% solution per reinforcer). The other outcome was provided later in their home cage and used as a control for general satiation in the outcome devaluation test. Before training commenced, mice were food restricted to 90% of their baseline weight; they were maintained at this lower weight for the duration of experimental procedures.

Training was conducted as follows. Each day, each mouse was trained in two separate operant chambers distinguished by contextual cues (i.e., black-and-white, vertical-striped, laminated paper in 3.2-mm-wide stripes on chamber walls or clear plexiglass chamber walls). Upon completion of training in one context, mice were immediately trained in the remaining context. For each mouse, the order of schedule exposure, lever position, and outcome obtained upon lever pressing were kept constant across contexts. However, the context, schedule order, lever position, and outcome earned were counterbalanced across mice. Each training session commenced with illumination of the house light and lever extension and ended following schedule completion or after 60 min, with the lever retracting and the house light turning off.

On the first day, mice were trained to approach the food magazine (no lever present) in each context on a random time (RT) schedule, with a reinforcer delivered on average every 60 s for a total of 15 min. Next, mice were trained in each context on a continuous reinforcement schedule (continuous ratio of reinforcement [CRF] or fixed ratio 1 [FR1]), where every lever press was reinforced, with the possible number of earned reinforcers increasing across training days (CRF5, CRF15, and CRF30). In the absence of any predictive cue signaling reward delivery, mice acquired lever-press behavior within 3 days. After acquiring lever-press behavior, mice were trained on RI and RR schedules of reinforcement (Adams, 1982; Derusso et al., 2010), with schedules differentiated by context and the session ending in each context after 15 reinforcers were earned or after 60 min had elapsed. Mice initially pressed under RI30 (on average, one reinforcer following the first press after an average of 30 s) and RR10 (on average, one reinforcer every 10 lever presses) schedules for 2 days, followed by 4 days of RI60 and RR20 training.

Outcome devaluation testing occurred across 2 consecutive days. In brief, on the valued day, mice had ad libitum access to the home-cage outcome for 1 hr before serial, brief, non-reinforced test sessions in the previous RI and RR training contexts. On the devalued day, mice were given 1 hr ad libitum access to the outcome previously earned by lever press and then underwent serial, non-reinforced test sessions in each training context. Prefeeding took place in a cage separate from the one in which mice were previously habituated, and the amount consumed was recorded. Order of context exposure during testing was the same as training exposure, with order of devaluation day counterbalanced across mice. Tests in each context were 5 min in duration.

A subset of mice was given an additional probe lever-press test following outcome devaluation testing. Food-restricted mice were pretreated (15 min) with CNO (n = 6) or saline (n = 6) and then placed in the first of the previously trained operant contexts (counterbalanced across RI and RR within each treatment group). Mice had access to the training lever, but the reinforcement schedule was a fixed ratio of 8 (every eighth lever press produced the outcome). Mice could earn unlimited reward within a 1 hr period. Lever presses, reward earned, and head entry behavior was analyzed.

Locomotor Activity in a Novel Cage

Control and hM4D mice were injected with CNO (1 mg/kg, 10 mL/kg) and placed in a novel polycarbonate cage for 60 min. Horizontal activity was detected as infrared beam crosses (1 inch spacing, 10 beams per cage) within 10 s bins using Opto M3 activity monitors (Columbus Instruments). Once the trial was over, mice were immediately returned to their home cage.

Slice Preparation and Electrophysiology

Mice were anesthetized with isoflurane and decapitated. Brains were removed and placed in ice-cold cutting solution containing: 194 mM

sucrose, 30 mM NaCl, 4.5 mM KCl, 26 mM $NaHCO_3$, 1.2 mM NaH_2PO_4, and 10 mM glucose. Coronal brain slices containing the OFC or DS, 250 μm thick, were obtained using a vibrating blade microtome (Leica VT1200S) and recovered in aerated artificial cerebrospinal fluid [ACSF] containing 124 mM NaCl, 4.5 mM KCl, 1 mM $MgCl_2$, 26 mM $NaHCO_3$, 1.2 mM NaH_2PO_4, 10 mM D-glucose, and 2 mM $CaCl_2$ at 32°C for 30 min. Slices were then placed at room temperature until experimental use. Whole-cell patch clamp recordings were performed between 30°C and 32°C ± 1°C (with control by an automatic temperature controller, Warner Instruments). Neurons in slices were visualized with an upright microscope using a 40× (0.8 numerical aperture) water immersion objective. Real-time images were displayed on a video monitor, which aided navigation and placement of recording pipettes. Then, 2–4 MΩ patch pipettes were pulled from borosilicate glass capillaries (1.5 mm outer diameter, 0.86 mm inner diameter; World Precision Instruments) and filled with internal solution. Two internal solutions were used. The K-based internal contained 126 mM K-gluconate, 4 mM KCl, 10 mM HEPES, 4 mM Mg-ATP, 0.3 mM Na-guanosine triphosphate (GTP), and 10 mM phosphocreatine. The Cs-based internal contained 150 mM CsCl, 10 mM HEPES, 2 mM $MgCl_2$, 3 mM Mg-ATP, 0.3 mM Na-GTP, and 0.2 mM 1,2-bis(*o*-aminophenoxy)ethane-*N*,*N*,*N*′,*N*′-tetraacetic acid (BAPTA)-4Cs. Recordings were made using a Multiclamp 700B amplifier (Molecular Devices). Membrane currents were filtered at 2 kHz, digitized using a Digidata 1322A at 10 kHz, displayed and saved using Clampex v.9.2, and analyzed with Clampfit v.9.2 (Molecular Devices). Statistical analysis was performed with GraphPad Prism 6 software (GraphPad Software). To isolate EPSCs, picrotoxin (100 μM; Sigma) was added to the extracellular solution. Series resistance was less than 25 MΩ, and cells with changes of more than 20% were excluded. A 10 min application of WIN55,212 (1 μM) was used to examine presence or absence of CB1-mediated synaptic depression. For optoactivation experiments, oEPSCs in MSNs and oEPSPs in OFC neurons were evoked in brain slices using 470 nm blue light (5 ms exposure time) delivered via field illumination using a high-power light-emitting diode (LED) source (LED4D067, Thor Labs). Light intensity was adjusted to produce oEPSCs with magnitude of 100–400 pA (<100 mW). The oEPSCs were evoked once per minute. Imaging of brain slices was performed using an Olympus MVX10 microscope (Olympus Corporation of America).

Statistical Analyses

The α level was set at 0.05 for all analyses unless indicated otherwise. Initial analyses showed normal distributions for all behavioral data. For all behavioral analyses, lever presses, lever-press rate, reward earned, and head entries were analyzed using repeated-measures ANOVA, with post hoc analyses performed using Bonferroni-corrected paired t tests where appropriate. For outcome devaluation testing analyses, two-way ANOVA (Devaluation state × Schedule) within each group was used to evaluate differences in lever-press and consumption behavior, with post hoc analyses performed using Bonferroni-corrected paired t tests where appropriate. To investigate the within-subject distribution of lever presses between valued and devalued states, we normalized lever presses for valued and devalued states to total lever pressing (Valued + Devalued) in each context. We then conducted planned one-sample t tests for normalized data to examine whether each condition differed from chance (0.5)—that is, the distribution of lever presses between valued and devalued states for each schedule observed in normalized data, with a value of 0.5 reflecting the same level of lever pressing between valued and devalued states. In addition, we examined the magnitude of outcome devaluation by creating a devaluation index [(Valued lever presses − Devalued lever presses)/Total lever presses] for each mouse in the RR and the RI contexts. We then conducted paired t tests to examine differences in the magnitude of devaluation between RI and RR contexts for each group of mice. For slice experiments, ten optically evoked EPSCs recorded before WIN administration were averaged to calculate the baseline amplitude, and ten oEPSCs recorded 10 min following the completion of drug application were averaged to determine the postdrug amplitude. Postdrug amplitudes were normalized to predrug amplitudes, and an unpaired Student's t test of the normalized amplitudes from wild-type and $CB1^{flox}$ animals was performed.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2016.04.043.

AUTHOR CONTRIBUTIONS

C.M.G., D.M.L., and R.M.C. designed the study, interpreted results, and wrote the manuscript. C.M.G. coordinated the experiments and performed surgical procedures and behavioral experiments. C.M.G., J.H.C., and B.K.A. performed in vitro electrophysiological experiments. C.M.G. and J.H.C. analyzed data. G.L. made mouse constructs. R.N., C.R., and K.D. supplied viruses.

ACKNOWLEDGMENTS

We thank Matthew Pava for assistance in the electrophysiological experiments and Emily Baltz for assistance with intracranial surgeries. We thank Drs. Beat Lutz and Giovanni Marsicano for providing the $Cnr1^{loxP/loxP}$ mice. The DREADD virus was provided by the Gene Therapy and Vector Core at the University of North Carolina. This research was supported by the NIAAA Division of Intramural Clinical and Biological Research, ERA-NET; European Research Council (COG 617142) and HHMI (IEC 55007415) grants to R.M.C.; and Pathway to Independence Award R00 AA021780 and a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation to C.M.G.

Received: October 16, 2015
Revised: March 16, 2016
Accepted: April 26, 2016
Published: May 26, 2016

REFERENCES

Adams, C.D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. Q. J. Exp. Psychol. B *34*, 77–98.

Adams, C.D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. Q. J. Exp. Psychol. B *33*, 109–121.

Ahmari, S.E., Spellman, T., Douglass, N.L., Kheirbek, M.A., Simpson, H.B., Deisseroth, K., Gordon, J.A., and Hen, R. (2013). Repeated cortico-striatal stimulation generates persistent OCD-like behavior. Science *340*, 1234–1239.

Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S., and Roth, B.L. (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.

Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.

Belin, D., and Everitt, B.J. (2008). Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. Neuron *57*, 432–441.

Belin, D., Belin-Rauscent, A., Murray, J.E., and Everitt, B.J. (2013). Addiction: failure of control over maladaptive incentive habits. Curr. Opin. Neurobiol. *23*, 564–572.

Bradfield, L.A., Dezfouli, A., van Holstein, M., Chieng, B., and Balleine, B.W. (2015). Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. Neuron *88*, 1268–1280.

Burguière, E., Monteiro, P., Feng, G., and Graybiel, A.M. (2013). Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. Science *340*, 1243–1246.

Burguière, E., Monteiro, P., Mallet, L., Feng, G., and Graybiel, A.M. (2015). Striatal circuits, habits, and implications for obsessive-compulsive disorder. Curr. Opin. Neurobiol. *30*, 59–65.

Colwill, R.M., and Rescorla, R.A. (1985). Postconditioning devaluation of a reinforcer affects instrumental responding. J. Exp. Psychol. Anim. Behav. Process. *11*, 120–132.

Corbit, L.H., Nie, H., and Janak, P.H. (2012). Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. Biol. Psychiatry *72*, 389–395.

DePoy, L., Daut, R., Brigman, J.L., MacPherson, K., Crowley, N., Gunduz-Cinar, O., Pickens, C.L., Cinar, R., Saksida, L.M., et al. (2013). Chronic alcohol produces neuroadaptations to prime dorsal striatal learning. Proc. Natl. Acad. Sci. USA *110*, 14783–14788.

Derusso, A.L., Fan, D., Gupta, J., Shelest, O., Costa, R.M., and Yin, H.H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. Front. Integr. Nuerosci. *4*, 17.

Dias-Ferreira, E., Sousa, J.C., Melo, I., Morgado, P., Mesquita, A.R., Cerqueira, J.J., Costa, R.M., and Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. Science *325*, 621–625.

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

Dickinson, A., Wood, N., and Smith, J.W. (2002). Alcohol seeking by rats: action or habit? Q. J. Exp. Psychol. B *55*, 331–348.

Fenno, L.E., Mattis, J., Ramakrishnan, C., Hyun, M., Lee, S.Y., He, M., Tucciarone, J., Selimbeyoglu, A., Berndt, A., Grosenick, L., et al. (2014). Targeting cells with single vectors using multiple-feature Boolean logic. Nat. Methods *11*, 763–772.

Gerdeman, G., and Lovinger, D.M. (2001). CB1 cannabinoid receptor inhibits synaptic release of glutamate in rat dorsolateral striatum. J. Neurophysiol. *85*, 468–471.

Gerdeman, G.L., Ronesi, J., and Lovinger, D.M. (2002). Postsynaptic endocannabinoid release is critical to long-term depression in the striatum. Nat. Neurosci. *5*, 446–451.

Gillan, C.M., Papmeyer, M., Morein-Zamir, S., Sahakian, B.J., Fineberg, N.A., Robbins, T.W., and de Wit, S. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. Am. J. Psychiatry *168*, 718–726.

Goldstein, R.Z., and Volkow, N.D. (2011). Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. Nat. Rev. Neurosci. *12*, 652–669.

Gourley, S.L., Olevska, A., Zimmermann, K.S., Ressler, K.J., Dileone, R.J., and Taylor, J.R. (2013). The orbitofrontal cortex regulates outcome-based decision-making via the lateral striatum. Eur. J. Neurosci. *38*, 2382–2388.

Graybiel, A.M. (2008). Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci. *31*, 359–387.

Gremel, C.M., and Costa, R.M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun. *4*, 2264.

Griffiths, K.R., Morris, R.W., and Balleine, B.W. (2014). Translational studies of goal-directed action as a framework for classifying deficits across psychiatric disorders. Front. Syst. Neurosci. *8*, 101.

Hilário, M.R.F., Clouse, E., Yin, H.H., and Costa, R.M. (2007). Endocannabinoid signaling is critical for habit formation. Front. Integr. Nuerosci. *1*, 6.

Hilario, M., Holloway, T., Jin, X., and Costa, R.M. (2012). Different dorsal striatum circuits mediate action discrimination and action generalization. Eur. J. Neurosci. *35*, 1105–1114.

Hoover, W.B., and Vertes, R.P. (2011). Projections of the medial orbital and ventral orbital cortex in the rat. J. Comp. Neurol. *519*, 3766–3801.

Kano, M., Ohno-Shosaku, T., Hashimotodani, Y., Uchigashima, M., and Watanabe, M. (2009). Endocannabinoid-mediated control of synaptic transmission. Physiol. Rev. *89*, 309–380.

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

Lovinger, D.M. (2010). Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. Neuropharmacology *58*, 951–961.

Marsicano, G., Goodenough, S., Monory, K., Hermann, H., Eder, M., Cannich, A., Azad, S.C., Cascio, M.G., Gutiérrez, S.O., van der Stelt, M., et al. (2003). CB1 cannabinoid receptors and on-demand defense against excitotoxicity. Science *302*, 84–88.

Mátyás, F., Yanovsky, Y., Mackie, K., Kelsch, W., Misgeld, U., and Freund, T.F. (2006). Subcellular localization of type 1 cannabinoid receptors in the rat basal ganglia. Neuroscience *137*, 337–361.

Nazzaro, C., Greco, B., Cerovic, M., Baxter, P., Rubino, T., Trusel, M., Parolaro, D., Tkatch, T., Benfenati, F., Pedarzani, P., and Tonini, R. (2012). SK channel modulation rescues striatal plasticity and control over habit in cannabinoid tolerance. Nat. Neurosci. *15*, 284–293.

Ostlund, S.B., and Balleine, B.W. (2007). The contribution of orbitofrontal cortex to action selection. Ann. N Y Acad. Sci. *1121*, 174–192.

Pan, W.X., Mao, T., and Dudman, J.T. (2010). Inputs to the dorsal striatum of the mouse reflect the parallel circuit architecture of the forebrain. Front. Neuroanat. *4*, 147.

Remijnse, P.L., Nielen, M.M.A., van Balkom, A.J.L.M., Cath, D.C., van Oppen, P., Uylings, H.B.M., and Veltman, D.J. (2006). Reduced orbitofrontal-striatal activity on a reversal learning task in obsessive-compulsive disorder. Arch. Gen. Psychiatry *63*, 1225–1236.

Rhodes, S.E.V., and Murray, E.A. (2013). Differential effects of amygdala, orbital prefrontal cortex, and prelimbic cortex lesions on goal-directed behavior in rhesus macaques. J. Neurosci. *33*, 3380–3389.

Rolls, E.T., Everitt, B.J., and Roberts, A. (1996). The orbitofrontal cortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. *351*, 1433–1443, discussion 1443–1444.

Rothermel, M., Brunert, D., Zabawa, C., Díaz-Quesada, M., and Wachowiak, M. (2013). Transgene expression in target-defined neuron populations mediated by retrograde infection with adeno-associated viral vectors. J. Neurosci. *33*, 15195–15206.

Samson, H.H., Cunningham, C.L., Czachowski, C.L., Chappell, A., Legg, B., and Shannon, E. (2004). Devaluation of ethanol reinforcement. Alcohol *32*, 203–212.

Schilman, E.A., Uylings, H.B.M., Galis-de Graaf, Y., Joel, D., and Groenewegen, H.J. (2008). The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. Neurosci. Lett. *432*, 40–45.

Schoenbaum, G., and Shaham, Y. (2008). The role of orbitofrontal cortex in drug addiction: a review of preclinical studies. Biol. Psychiatry *63*, 256–262.

Shepherd, G.M.G. (2013). Corticostriatal connectivity and its role in disease. Nat. Rev. Neurosci. *14*, 278–291.

Smith, K.S., and Graybiel, A.M. (2013). A dual operator view of habitual behavior reflecting cortical and striatal dynamics. Neuron *79*, 361–374.

Stachniak, T.J., Ghosh, A., and Sternson, S.M. (2014). Chemogenetic synaptic silencing of neural circuits localizes a hypothalamus→midbrain pathway for feeding behavior. Neuron *82*, 797–808.

Stalnaker, T.A., Calhoon, G.G., Ogawa, M., Roesch, M.R., and Schoenbaum, G. (2010). Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Front. Integr. Nuerosci. *4*, 12.

Stalnaker, T.A., Cooch, N.K., and Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. Nat. Neurosci. *18*, 620–627.

Thorn, C.A., Atallah, H., Howe, M., and Graybiel, A.M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. Neuron *66*, 781–795.

Uchigashima, M., Narushima, M., Fukaya, M., Katona, I., Kano, M., and Watanabe, M. (2007). Subcellular arrangement of molecules for 2-arachidonoyl-glycerol-mediated retrograde signaling and its physiological contribution to synaptic modulation in the striatum. J. Neurosci. *27*, 3663–3676.

Wall, N.R., De La Parra, M., Callaway, E.M., and Kreitzer, A.C. (2013). Differential innervation of direct- and indirect-pathway striatal projection neurons. Neuron *79*, 347–360.

Yin, H.H., and Knowlton, B.J. (2006). The role of the basal ganglia in habit formation. Nat. Rev. Neurosci. *7*, 464–476.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur. J. Neurosci. *19*, 181–189.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2005a). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. Eur. J. Neurosci. *22*, 505–512.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., and Balleine, B.W. (2005b). The role of the dorsomedial striatum in instrumental conditioning. Eur. J. Neurosci. *22*, 513–523.

Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav. Brain Res. *166*, 189–196.

Znamenskiy, P., and Zador, A.M. (2013). Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. Nature *497*, 482–485.