



Published in final edited form as:

*Eur J Neurosci.* 2013 August ; 38(3): . doi:10.1111/ejn.12239.

## The orbitofrontal cortex regulates outcome-based decision-making via the lateral striatum

Shannon L. Gourley<sup>1,2,3</sup>, Anastasia Olevska<sup>3</sup>, Kelsey S. Zimmermann<sup>2</sup>, Kerry J. Ressler<sup>2,4,5</sup>, Ralph J. DiLeone<sup>3,6</sup>, and Jane R. Taylor<sup>3,6,7</sup>

<sup>1</sup>Department of Pediatrics, Emory University

<sup>2</sup>Neuroscience Program, Emory University

<sup>3</sup>Department of Psychiatry, Division of Molecular Psychiatry, Yale University

<sup>4</sup>Department of Psychiatry and Behavioral Sciences, Emory University

<sup>5</sup>Howard Hughes Medical Institute

<sup>6</sup>Interdepartmental Neuroscience Program, Yale University

<sup>7</sup>Department of Psychology, Yale University

### Abstract

The orbitofrontal cortex (oPFC) sends substantial projections to the ventrolateral striatum and aspects of the nucleus accumbens that are—functionally—poorly understood. This is despite probable cortico-striatal involvement in multiple diseases such as addiction and obsessive-compulsive disorder. Here we surgically disconnected the oPFC from the ventrolateral striatum using unilateral asymmetric lesions in mice and classified instrumental decision-making strategies. Mice with symmetric lesions that spared one oPFC-striatal network served as controls. As a complementary approach, we selectively knocked down *Brain-derived neurotrophic factor (Bdnf)* bilaterally in the oPFC and ascertained behavioral and neurobiological consequences within the downstream striatum. oPFC-striatal disconnection and oPFC *Bdnf* knockdown blocked sensitivity to outcome-predictive relationships in both food-reinforced and cocaine-associated settings. *Bdnf* knockdown simultaneously regulated striatal BDNF expression, and striatal c-Fos predicted sensitivity to action-outcome associative contingencies. Prior evidence strongly implicates the *dorsolateral* striatum in stimulus-response habit formation. Our findings thus provide novel evidence for functional compartmentalization within the lateral striatum, with the dorsal compartment subserving classical stimulus-response habit systems and a ventral compartment coordinating outcome-based decision-making via oPFC interactions. This compartmentalization may apply to both ‘natural’—as in the case of food-reinforced behavior—and ‘pathological’—as in the case of cocaine-seeking—contexts.

### Keywords

orbital; habit; contingency; action; outcome; addiction

## Introduction

Considerable evidence indicates that both humans and rodents can associate specific actions with their outcomes, but that with repeated performance, familiar actions assume automated, stimulus-elicited, habitual qualities that are resistant to change. Converging neuroanatomical models characterize this process as a transition from dorsomedial and dorsolateral striatal systems (**DMS** and **DLS**) that act in concert, to a DLS-centric circuit (Yin et al., 2008, 2009; Kimchi et al., 2009). The DMS likely coordinates goal-directed behavior via connectivity with the medial prefrontal cortex, while the DLS coordinates stimulus-response habits via interactions with the sensorimotor cortices. Landmark lesion and inactivation studies targeted the dorsal-most aspects of the DLS (Yin et al., 2004, 2006), while the cell populations within the ventrolateral striatum have received relatively little attention, and their functional importance in decision-making remains unknown.

This gap in current knowledge is particularly notable because the orbitofrontal prefrontal cortex (oPFC) sends substantial topographically-organized projections to aspects of both the caudate/putamen and nucleus accumbens in primates and rodents (Schilman et al., 2008). Argued to play a prominent role in complex decision-making, as well as sensitivity to drugs of abuse (Rushworth et al., 2007; Schoenbaum et al., 2009; Lucantonio et al., 2012), the lateral oPFC projects to the ventrolateral striatum (Schilman et al., 2008). Whether oPFC interactions with the ventrolateral striatum *specifically* contribute to outcome-based decision-making has not to our knowledge been tested.

Striatal-targeted oPFC projections are organized ipsilaterally in the brain. This segregated anatomy allows for classical disconnection experiments in which a lesion is placed in one structure in each hemisphere. When lesions are asymmetric, one structure in each hemisphere remains intact, but the oPFC and ventrolateral striatum are “disconnected.” The benefit of this approach, when combined with the appropriate symmetric lesion control group, is that it allows for the assessment of the impact of oPFC-striatal interactions on instrumental decision-making. As a complementary approach here, we also knocked down *Bdnf* bilaterally in the oPFC and ascertained both behavioral and neurobiological consequences, namely c-Fos and BDNF expression, in the downstream striatum.

This report provides novel evidence for functional compartmentalization within the lateral striatum: While the dorsal component is strongly associated with classical stimulus-response habit systems (Yin et al., 2008), oPFC interactions with the ventrolateral compartment may by contrast guide outcome-based decision-making.

## Materials and Methods

### Subjects

Male C57BL/6 mice (Charles River Laboratories, Kingston, NY) and transgenic mice bred in-house and described below were maintained on a 12-hour light cycle (0700 on), experimentally naïve, and at least 10 weeks of age. Animals were provided food and water *ad libitum* except during instrumental conditioning when body weights were maintained at ~90% of baseline. Procedures were approved by the Yale and Emory University IACUCs and were compliant with National Institutes of Health guidelines regarding the care and use of animals for experimental procedures.

### *Bdnf* knockdown and lesion

Wild type mice or mice homozygous for a floxed allele (exon 5) encoding the *Bdnf* gene (Rios et al., 2001) were anaesthetized with 1:1 2-methyl-2-butanol and tribromoethanol (Sigma) diluted 40-fold with saline in the case of lesions or ketamine/xylazine in the case of

viral vector infusion. With needles centered at bregma, stereotaxic coordinates were located on the leveled skull (David Kopf Instruments, Tujunga, CA). AAV-EGFP+Cre recombinase was infused in a volume of 0.5  $\mu$ l at AP+2.6, DV-2.8, ML $\pm$ 1.2 (Bissonette et al., 2008; Gourley et al., 2010) over 5 min with needles left in place for 4 additional min. Mice were sutured and allowed to recover for at least 3 weeks, allowing for *Bdnf* knockdown.

Alternatively, NMDA (20  $\mu$ g/ $\mu$ l) was infused ipsilaterally at the same coordinates in a volume of 0.1  $\mu$ l over 1 min. with the needles left in place for an additional 2 min. NMDA infusions were then placed either ipsilaterally or contralaterally in the ventrolateral striatum at AP+0.5, DV-3.5, ML $\pm$ 2.7. Control mice were infused with saline—half were infused ipsilaterally, while half were infused contralaterally. Throughout, no differences were observed between these groups, or between saline-infused mice and mice infused with AAV-EGFP bilaterally into the oPFC. These three control groups were thus combined for statistical and graphical purposes.

### **Instrumental conditioning and action-outcome contingency degradation**

Mice were trained to nose poke for food reinforcement (20 mg grain-based pellets; Bioserv, Frenchtown, NJ) using standard illuminated Med-Associates (Georgia, VT) conditioning chambers with 3 nose poke recesses. Training was initiated with a continuous reinforcement schedule; 30 pellets were available for responding on the 2 outermost recesses, resulting in a maximum of 60 pellets/session. Sessions ended when all 60 pellets were delivered or at 135 min. Five daily sessions were conducted, during which animals acquired the response. Next, mice were shifted to a random interval (RI) 30-second schedule of reinforcement for 2 sessions; again, 30 pellets were available for responding on each of 2 apertures.

Action-outcome contingency degradation was accomplished over two sessions, the order of which was counter-balanced across mice. In the non-degraded session, one nose poke aperture was occluded, and responding on the other aperture was reinforced on a variable ratio 2 (VR2) schedule; mice were required to retrieve each pellet before earning more during the 25 min session. In the degraded session, the opposite aperture was occluded, and reinforcers were delivered into the magazine at a rate matched to each animal's reinforcement rate the previous day. Responses produced no programmed consequences (Gourley et al., 2012a; Hammond, 1980). Response rates during these two sessions were analyzed by 2-factor (aperture x group) analysis of variance (ANOVA).

### **Instrumental reversal**

We next tested mice in an instrumental reversal test in which mice were required to shift responding to a previously unreinforced center nose poke aperture. Responding on the outermost apertures (*i.e.*, the previously active apertures) was un-reinforced. A VR2 schedule of reinforcement was used, and mice were required to retrieve each pellet before acquiring another. Sessions were 25 min. long and conducted daily for 4 days. Response rates on the active aperture were analyzed by 2-factor (lesion x session) ANOVA with repeated measures. Response acquisition in this task is impaired by bilateral oPFC lesions targeting the lateral compartment, as here (Gourley et al., 2010).

### **Outcome devaluation**

We also tested sensitivity to devaluation of the food outcome using a satiety-specific prefeeding procedure. Here, trained mice were allowed 30 min. access to the reinforcer pellets in a clean cage before a 15-min. test session conducted in extinction. Responding was normalized to a non-devalued, *i.e.*, “valued,” session—a 15-min. test session also conducted in extinction, prior to which food pellets were *not* available. Devalued and “valued” sessions

were counter-balanced, and response rates were compared by ANOVA. There were no group differences in food intake during the prefeeding period (not shown).

### Cocaine-conditioned place preference

The conditioned place preference apparatus (Med-Associates) consisted of two distinct chambers (black with rod floor, white with mesh floor and dim lighting) connected by a central neutral chamber with vertical sliding doors. Photobeams recorded time spent in each compartment. Mice received a 20-min. pre-training test in which they had free access to all three chambers to evaluate individual place preference. Next, cocaine conditioning took place over 4 consecutive days during which each animal received cocaine immediately prior to being placed in its *non-preferred* chamber [2 sessions; 10 mg/kg *i.p.* cocaine-hydrochloride (Sigma) dissolved in saline], or a saline injection was paired with the preferred chamber (2 sessions). Mice remained confined to the respective chambers for 30 min. after injection.

Place preference was tested by placing mice in the middle neutral compartment with both doors open. Mice were allowed to explore freely for 20 min.—this is annotated in text as the “early test.” With 2 more cocaine-context pairings using 30 mg/kg cocaine, all mice developed a pronounced preference for the cocaine-associated chamber, as ascertained in an additional place preference test referred to as the “late test”. Next, the injection was withheld, but “context pairing” sessions were otherwise identical to those used during cocaine-context conditioning. Lastly, a final preference test was conducted. Throughout, place preference was calculated as time spent in the cocaine-paired chamber minus time spent in the other chambers and analyzed by ANOVA with repeated measures.

### Histology

Fixed tissue was soaked in 4% PFA for 48 hours, then transferred to 30% sucrose. Tissues were sliced into 40- $\mu$ m thick sections on a microtome held at  $-15^{\circ}\text{C}$ . In the case of viral vector delivery, every third slice through the prefrontal cortex was mounted, and EGFP was imaged. To confirm lesion sites, every third section was immunostained for NeuN (Millipore, Billerica, MA; Rb; 1:500) and Glial Fibrillary Acidic Protein (GFAP) (Dakocytomation, Carpinteria, CA; Ms; 1:1000). AlexaFluor goat IgGs (Invitrogen, Carlsbad, CA; 1:300) served as secondary antibodies. EGFP and GFAP signals were graphically transposed onto corresponding images from the mouse brain atlas (Paxinos and Franklin, 2003). In figure 1, gray represents the largest lesion or virus spread, while black is the smallest. Lesions are shown as a red stain (GFAP), while green signals EGFP or NeuN as indicated.

### BDNF quantification and immunoblotting

*Bdnf* knockdown experiments were conducted in 2 cohorts: In the first cohort, mice were transcardially perfused with 4% paraformaldehyde (PFA) after deep sedation with pentobarbital, and brains were extracted for histological analysis as described above. In the second cohort, mice were rapidly decapitated after the last session, and fresh brains were transected—the rostral component was submerged in 4% PFA for 48 hours for identical histological analyses, and the caudal component was immediately frozen on dry ice for subsequent immunoblotting and BDNF quantification by enzyme-linked immunosorbent assay (ELISA). Frozen tissue was sliced into 1 mm-thick coronal sections, and the lateral striatum dorsal to the nucleus accumbens and amygdala were extracted with bilateral tissue punches (1 mm core) and homogenized in lysis buffer [200  $\mu$ l: 137 mM NaCl, 20 mM Tris-HCl (pH=8), 1% Igepal, 10% glycerol, 1:100 Phosphatase Inhibitor Cocktails 1 and 2 (Sigma)] by sonication.

Western blotting and ELISA were conducted as previously described (Gourley et al., 2008, 2009): In brief, for immunoblotting, 20  $\mu$ g/sample were added to 10  $\mu$ l Laemmli buffer (20% glycerol, 2% SDS, Bromphenol blue) and boiled for 10 min. Samples were separated by SDS-PAGE on 8–16% gradient tris-glycine gels (Invitrogen, Carlsbad, CA). Primary antibodies were anti-GAPDH (Ms; 1:20K; Advanced Immunochemical Inc., Long Beach, CA) and anti-c-Fos (Rb; 1:500; Santa Cruz Biotechnology, Santa Cruz, CA). For antibody detection by the Odyssey infrared imaging system (LI-COR, Lincoln, NE), membranes were incubated with IRDye 700 Dx Anti-Rb IgG and IRDye 800 Dx Anti-Ms IgG (1:5,000; Rockland Immunochemicals, Gilbertsville, PA). Individual bands were quantified using Odyssey software.

BDNF quantification was conducted using a 2-site BDNF ELISA kit in accordance with the manufacturer's instructions (Promega, Madison, WI), except tissue was diluted 1:1, and the extraction procedure was excluded. Samples were run in duplicate, and BDNF concentrations were normalized to the total protein content in each sample. BDNF was analyzed by ANOVA.

### Statistical analyses

Two-tailed parametric statistical analyses with  $\alpha = 0.05$  were performed using SigmaStat v. 3.1 and SPSS. Dependent measures were analyzed by 2- or 3-factor ANOVA as appropriate, and Tukey's post-hoc tests were utilized in the event of significant interaction effects; significant post-hoc effects are indicated graphically. In some cases, only main effects were detected; these instances are indicated textually and graphically by an asterisk in the corresponding legend. For c-Fos analysis, infrared signal values generated by LiCor densitometry analysis were correlated using Spearman's correlation co-efficient with response rates on the active apertures during the final day of instrumental conditioning.

### Results

Prefrontal cortical lesions and viral infections were largely located within the lateral subregion of the oPFC; in some mice, lesions spread to the ventral compartment (fig. 1a). Dorsal striatal lesions affected the lateral compartment of the dorsal striatum in all animals, and some spread to the intermediate dorsal striatum was noted. In accordance with known projection patterns, our lesions targeted the ventral compartment of the lateral striatum (though some spread to the DLS was noted), and spared the nucleus accumbens.

When mice were trained to nose-peek for food reinforcers, response rates in the knockdown and disconnection groups lagged relative to saline-infused control mice and those with ipsilateral lesions sparing one oPFC-ventrolateral striatal network. This lag was particularly apparent when the reinforcement schedule escalated from a fixed ratio to a random interval in the final 2 training sessions [group x session interaction  $F_{(18,198)}=2.6, p<0.001$ ] (fig. 1b).

This profile has been previously associated with an impairment in action-outcome associative conditioning (Corbit and Balleine, 2003), hence we next degraded the action-outcome relationship associated with one of the two active apertures by providing food reinforcers non-contingently ("contingency degradation"). In this case, mice with oPFC *Bdnf* knockdown and oPFC-ventrolateral striatum disconnection failed to differentiate between the degraded and non-degraded apertures [group x aperture interaction  $F_{(3,33)}=4.6, p=0.009$ ] (fig. 1c). A comparison between fig. 1b and 1c makes clear that these animals responded on both the degraded and nondegraded apertures at the same rate as during instrumental conditioning when both apertures were reinforced.

We next utilized an instrumental reversal task sensitive to oPFC lesions to further characterize decision-making strategies. In this case, mice were required to respond on a center, previously inactive aperture. This task does not use cues signaling either reinforcement delivery or a failure to perform a reinforced response, and thus requires the animal to acquire a new action-outcome, as opposed to stimulus-response or stimulus-outcome, association. Here, mice with both *Bdnf* knockdown and oPFC-ventrolateral striatum disconnection were again impaired. Restricting our analysis first to responding on the reinforced aperture, a main effect of group indicated that mice with contralateral lesions and oPFC *Bdnf* knockdown responded less than control counterparts [main effect of group  $F_{(3,31)}=2.9, p<0.05$ ] (fig. 1d).

More critically, when responding on the previously reinforced aperture was also considered in our analysis, a 3-way group x aperture x session interaction was identified [interaction  $F_{(3,62)}=3.6, p=0.01$ ] (fig. 1e, left panel). Post-hoc comparisons can be summarized as follows (see also fig. 1e, right panel): During the first of four extinction training sessions, all mice regardless of group responded preferentially on the previously reinforced aperture. By the fourth session, however, control mice and mice with ipsilateral lesions preferentially responded on the newly reinforced aperture and by comparison neglected the previously reinforced aperture. In other words, these mice showed successful instrumental reversal and extinction of a nonreinforced response. By contrast, mice with contralateral lesions or oPFC *Bdnf* knockdown responded non-selectively on both apertures even after 4 training sessions. This pattern of responding echoes that observed during action-outcome contingency degradation, in which these experimental mice also failed to refine their response strategies in response to changes in the action-outcome associative contingency.

We next measured BDNF concentrations in the downstream striatum and amygdala in tissues collected from mice with oPFC *Bdnf* knockdown and found diminished BDNF expression levels [main effect of knockdown  $F_{(1,10)}=9.4, p=0.01$ , no interactions] (fig. 1f). Moreover, expression of the immediate early gene c-Fos in the striatum correlated with goal-directed decision-making [ $r=0.48, p<0.05$ ]. Specifically, response rates on the active aperture were associated with high striatal c-Fos expression levels, while low response rates were associated with low c-Fos (fig. 1g).

Despite these deficiencies, *Bdnf* knockdown and oPFC-ventrolateral striatum disconnection notably spared sensitivity to reductions in the value of the food outcome. Here, satiety-specific outcome devaluation reduced instrumental responding in all groups regardless of lesion or *Bdnf* status (main effect of devalued outcome  $p<0.001$ ; interaction between devalued outcome and group  $F<1$ ) (fig. 1h).

A failure to engage in goal-directed outcome-based decision-making, relative to stimulus-based habits, is thought to play a causal role in addiction (Jentsch and Taylor, 1999; Everitt and Robbins, 2005). Thus, as a final experiment, we generated mice with oPFC-selective *Bdnf* deficiency and utilized a conditioned place preference model of drug-seeking, pairing cocaine and saline with unique test environments. A probe test after only two cocaine pairings indicated that mice with *Bdnf* knockdown—but not control mice—had already developed cocaine-conditioned place preference [group x test interaction (including all probe tests)  $F_{(3,31)}=5.9, p=0.002$ ] (fig. 2). In other words, BDNF-deficient mice developed cocaine-conditioned place preference more rapidly than control counterparts. With multiple pairings, all mice ultimately developed cocaine-conditioned place preference (fig. 2). Importantly, when the predictive relationship between the cocaine-associated chamber and the cocaine was degraded by withholding cocaine, place preference returned to chance levels in control mice almost immediately, while preference was unremitting, or ‘habitual,’ in knockdown mice (fig. 2).



## Discussion

It is now widely appreciated that instrumental decision-making can be informed by the predictive relationship between an action and its outcome, or by the value of the outcome (for review, Yin et al., 2008). Homologous cortico-striatal networks in humans and rodents are thought to regulate response selection and choice behavior (Balleine and O'Doherty, 2010); hence, a firmer grasp on the neuroanatomical and neurobiological mechanisms of outcome-based decision-making in rodent models is a fundamental research imperative that may shed light onto diseases characterized by aberrant or deficient sensitivity to action-outcome associative contingencies. Here we generated mice in which the oPFC was disconnected from the downstream ventrolateral striatum by virtue of unilateral asymmetric lesions. When mice were trained to nose poke for food reinforcers, response rates in these mice lagged relative to control mice, as well as those with lesions that were restricted to a single hemisphere. This pattern has been associated with impaired action-outcome associative learning (Corbit and Balleine, 2003). Consistent with this perspective, mice with oPFC-ventrolateral striatum disconnections failed to differentiate between a 'degraded' and 'non-degraded' instrumental response after action-outcome contingency degradation. Moreover, both oPFC-selective *Bdnf* knockdown and oPFC-ventrolateral striatum disconnection impaired response refinement in an instrumental reversal task that is sensitive to oPFC lesions (Gourley et al., 2010). Specifically, these mice showed impaired response acquisition and ultimately failed to differentiate between the nonreinforced, extinguished apertures relative to the newly reinforced aperture. Together, this pattern suggests that oPFC innervation of the ventrolateral striatum coordinates outcome-based decision-making, and that BDNF is a critical molecular determinant of oPFC function.

What is the nature of oPFC innervation of the striatum? In a comprehensive analysis of corticostriatal projections from the rodent oPFC, Groenewegen and colleagues reported that the oPFC projects in a topographically-organized fashion to the striatum (Schilman et al., 2008). For example, the medial orbital cortex (lying ventral to the prelimbic cortex in the frontal pole) innervates the medial-most DMS, and the ventral oPFC projects by contrast to a vertical column of cells within the intermediate dorsal striatum. The lateral and dorsolateral oPFC innervate the ventrolateral striatal compartments, targeting subregions both within, and dorsal to, the nucleus accumbens. This pattern is in agreement with earlier studies (Berendse et al., 1992; Reep et al., 1996, 2003) that provided initial evidence that corticostriatal projections are organized similarly between rodents and non-human primates; for example, there is considerable overlap between orbital and medial prefrontal terminals in the DMS in both species. One difference, however, appears to be the patterns of innervation of the nucleus accumbens core, which is largely devoid of oPFC projections in the rodent, but not in the primate (discussed Schilman et al., 2008).

Our lesions here targeted the ventrolateral compartment of the oPFC and the corresponding ventrolateral striatum and when placed in contralateral hemispheres, impaired action-outcome associative conditioning in a contingency degradation paradigm. By contrast, pioneering studies by Yin, Knowlton, and Balleine showed that rats with bilateral *DLS* lesions or local inactivation develop outcome-based goal-directed instrumental response strategies and in fact, are *resistant* to the development of stimulus-response habits (Yin et al., 2004, 2006). This pattern – the opposite of that reported here – strikingly dissociates subregions of the lateral striatum: Specifically, we propose a model wherein the dorsal aspect, notably devoid of oPFC innervation (Schilman et al., 2008), regulates stimulus-response habit formation via interactions with sensorimotor cortices (Yin et al., 2008). By contrast, the ventral compartment, innervated by the lateral oPFC, coordinates action-outcome associative conditioning (fig. 1i).

Despite their behavioral deficiencies, mice with oPFC-ventrolateral striatum disconnection retained sensitivity to reductions in the value of the food outcome, in that satiety-specific outcome devaluation reduced instrumental responding. At the same time, sensitivity to devaluation was qualitatively less pronounced in our comparison group of mice with bilateral oPFC *Bdnf* knockdown. How might we interpret these findings? First, it is notable that oPFC *Bdnf* knockdown deprives both the striatum and amygdala of BDNF protein (fig. 1f), and previous work indicates that the basolateral amygdala regulates sensitivity to outcome value in tasks in which action-outcome decision-making is measured by altering the value of the outcome (Balleine et al., 2003). Together with substantial evidence that BDNF plays a permissive role in activity-dependent long-term plasticity within the basolateral amygdala (e.g., see Li et al., 2011), we posit that deprivation of the basolateral amygdala of BDNF resulted in a modest (non-significant) deficit in sensitivity to outcome value in the *Bdnf* knockdown group, but that an oPFC-ventrolateral striatal circuit is a relatively selective determinant of action selection when the action is associated with a given outcome, though not when that action is dependent on the value of the outcome. This is notably distinct from sensitivity to reward prediction error or to outcomes that are signaled by environmental stimuli; in these cases, the oPFC has been implicated in behavioral decision-making (Rushworth et al., 2007; Sul et al., 2010).

### Orbital BDNF systems as determinants of decision-making

A primary finding of this report is that oPFC *Bdnf* knockdown restricted to the oPFC impaired outcome-based decision-making under both food-reinforced and cocaine-associated conditions. How might oPFC BDNF be acting? First, cortical *Bdnf* knockdown could impact decision-making strategies by retarding anterograde BDNF transport to, or BDNF synthesis in, major projection sites (Gourley et al., 2009; further discussed Choi et al., 2012). When we measured BDNF concentrations in the downstream lateral striatum and amygdala after gene knockdown within the oPFC, we indeed found diminished BDNF expression levels. Moreover, expression of the immediate early gene c-Fos in the lateral striatum correlated with goal-directed decision-making such that high DLS c-Fos expression levels predicted high response rates on the active apertures, while low response rates were associated with low c-Fos. Although neuroplasticity within the lateral striatum is more classically associated with stimulus-response habits (Yin et al., 2008), these findings are, as discussed in the previous section, consistent with known projection patterns from prefrontal cortical regions, including the oPFC to the ventrolateral striatum (Shilman et al., 2008), as well as with high rates of neuronal firing in the DLS during the initiation of goal-directed response selection in food-reinforced rats (Kimchi et al., 2009).

Drug-seeking behavior is commonly codified as habitual, stimulus-elicited behavior that is resistant to modification, and the val66met *bdnf* gene variant is associated with addiction vulnerability in humans and diminished activity-dependent BDNF secretion (Cheng et al., 2005; Egan et al., 2003). Therefore, we also utilized a conditioned place preference model of drug-seeking and paired cocaine and saline with unique test environments in control and *Bdnf* knockdown mice. *Bdnf* knockdown mice developed cocaine-conditioned place preference more rapidly, but more relevant to this report, when the predictive relationship between the cocaine-associated chamber and the cocaine was degraded by withholding cocaine, place preference was unremittingly stimulus-dependent, or 'habitual,' in knockdown mice. In this case, the associative structure differed from our experiments in fig. 1 in the sense that a stimulus – the cocaine-paired chamber – rather than an action predicted a certain outcome; nonetheless, these findings are consistent with a prior report that bilateral oPFC lesions impair decision-making that is based on stimulus-outcome associative relationships (Ostlund and Balleine, 2007), and they extend these findings by pointing to BDNF as a critical molecular determinant.



Notably, there are now multiple reports of intact sensitivity to outcome devaluation in situations in which rodents with oPFC lesions must associate a specific action (as opposed to a stimulus) with an outcome (see for example fig. 1 here; Ostlund and Balleine, 2007; Gourley et al., 2010). Thus, while we believe we have provided strong evidence that the oPFC is a critical determinant of action-outcome associative learning, it is important to note that its role is restricted relative to, for example, that of the prelimbic prefrontal cortex or the DMS. Lesions of *these* structures occlude sensitivity to modifications of outcome value, *as well as* the predictive value of the action, in tasks in which the animal must associate a specific action with a specific outcome (Balleine and Dickinson, 1998; Killcross and Coutureau, 2003; Corbit and Balleine, 2003; Yin et al., 2005). Thus, while the oPFC is clearly involved in outcome-based decision-making (fig. 1 here; reviewed Lucantonio et al., 2012), its role is distinct from that of the medial wall prefrontal cortical structures.

A final caveat is that the organization of oPFC projections to the medial and lateral striatum likely results in distinct functions of discrete subregions of the oPFC. While we propose a role for the lateral oPFC in outcomes-based decision-making here that is somewhat redundant with that the role of the medial orbitofrontal cortex that we previously described (Gourley et al., 2010), early evidence indicates the *ventral* oPFC is functionally distinct. Specifically, *Bdnf* knockdown in this region, which innervates a discrete vertical column of cells within the dorsal intermediate striatum (Schilman et al., 2008), *increases* instrumental responding for food reinforcers in an un-cued instrumental conditioning task (DePoy et al., 2013). Aberrant responding was normalized with pharmacological interventions targeting the actin cytoskeleton, leading us to posit that BDNF-mediated modifications in dendritic spine morphology and function resulted in increased sensitivity to non-discrete stimuli – such as the experimenter, the conditioning chamber, *etc.* – which then enhanced instrumental responding. Although further research is necessary, these findings nonetheless suggest that the ventral oPFC is functionally distinct from the lateral oPFC.

## Concluding Remarks

Previous investigations from our and others' labs using instrumental conditioning approaches to dissecting decision-making strategies indicate that the oPFC orchestrates outcome-based decision-making by regulating sensitivity to: *1*) action-outcome and stimulus-outcome predictive relationships (Gourley et al., 2010; Ostlund and Balleine, 2007; Lucantonio et al., 2012), and *2*) by encoding and/or updating the value of expected outcomes that are signaled by discrete conditioned stimuli (see for example Rushworth et al., 2007). Our findings are in agreement with the perspective that the oPFC updates action-outcome associations (see Sul et al., 2010), though not outcome value (see Takahashi et al., 2009). We also provide novel evidence that oPFC BDNF is a critical substrate of oPFC-dependent decision-making. Based on these findings, we propose a model wherein BDNF supports neuroplasticity at cortico-cortical and cortico-striatal synapses critical to outcome-guided decision-making (Thoenen 1995; Shen and Cowen, 2010). Moreover, given that oPFC dendritic spine stability is required for stimulus-outcome learning (Gourley et al., 2012b), BDNF may also act via spine-stabilizing influences in mature oPFC neurons (Vigers et al., 2012), which both receive subcortical projections from, and in turn synapse onto, lateral striatal medium spiny neurons (Thomas et al., 2000). In this model, BDNF is a critical regulator of outcome-based decision-making in both 'natural' contexts—as in the case of food-reinforced behavior—and 'pathological' contexts—as in the case of cocaine-associated decision-making.

## Acknowledgments

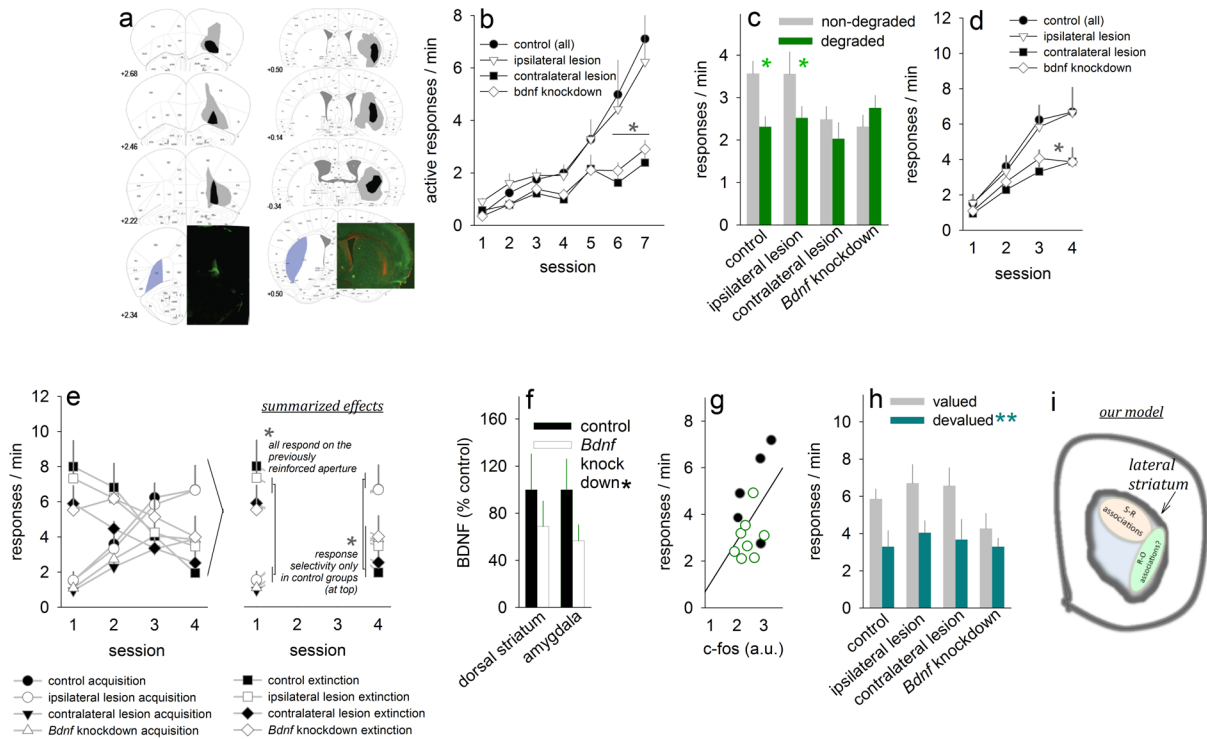
The authors thank Dr. Dennis C. Choi and colleagues in the Taylor laboratory for advice and feedback, and Dr. Daeyeol Lee for helpful suggestions on the manuscript. This work was supported by DA011717, UL1-DE19586

and the Roadmap for Medical Research/Common Fund, AA017537; the Connecticut Mental Health Center; and Children's Healthcare of Atlanta. The components of this project performed at the Yerkes National Primate Research Center were also funded by the National Center for Research Resources P51RR165 (currently Office of Research Infrastructure Programs/OD P51OD11132). The authors report no conflicts of interest.

## References

- Ballaine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 1998; 37:407–419. [PubMed: 9704982]
- Balleine BW, Killcross AS, Dickinson A. The effect of lesions of basolateral amygdala on instrumental conditioning. *J Neurosci*. 2003; 23:666–675. [PubMed: 12533626]
- Balleine BW, O'Doherty JP. Human and rodent homologues in action control: Corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010; 35:48–69. [PubMed: 19776734]
- Berendse HW, Galis-de Graaf Y, Groenewegen HJ. Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *J Comp Neurol*. 1992; 316:314–317. [PubMed: 1577988]
- Bissonette GB, Martins GJ, Franz TM, Harper ES, Schoenbaum G, Powell EM. Double dissociation of the effects of medial and orbital prefrontal cortical lesions on attentional and affective shifts in mice. *J Neurosci*. 2008; 28:11124–11130. [PubMed: 18971455]
- Corbit LH, Balleine BW. The role of the prelimbic cortex in instrumental conditioning. *Behav Brain Res*. 2003; 146:145–157. [PubMed: 14643467]
- Cheng CY, Hong CJ, Yu YW, Chen TJ, Wu HC, Tsai SJ. Brain-derived neurotrophic factor (Val66Met) genetic polymorphism is associated with substance abuse in males. *Brain Res Mol Brain Res*. 2005; 140:86–90. [PubMed: 16109452]
- Choi DC, Gourley SL, Ressler KJ. Prelimbic BDNF and trkB signaling regulates consolidation of both appetitive and aversive emotional learning. *Trans Psychiatry*. 2012; 2:e205.
- DePoy LM, Noble B, Allen AG, Gourley SL. Developmentally divergent effects of Rho-kinase inhibition on cocaine- and BDNF-induced behavioral plasticity. *Behav Brain Res*. 2013; 243:171–175. [PubMed: 23327740]
- Egan MF, Kojima M, Callicott JH, Goldberg TE, Kolachana BS, Bertolino A, Zaitsev E, Gold B, Goldman D, Dean M, Lu B, Weinberger DR. The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*. 2003; 112:257–269. [PubMed: 12553913]
- Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*. 2005; 8:1481–1489. [PubMed: 16251991]
- Gourley SL, Kiraly DD, Howell JL, Olausson P, Taylor JR. Acute hippocampal BDNF restores motivational and forced swim performance after corticosterone. *Biol Psychiatry*. 2008; 664:884–890. [PubMed: 18675955]
- Gourley SL, Howell JL, Rios M, DiLeone RJ, Taylor JR. Prelimbic cortex bdnf knock-down reduces instrumental responding in extinction. *Learn Mem*. 2009; 16:755–760.
- Gourley SL, Lee AS, Howell JL, Pittenger C, Taylor JR. Dissociable regulation of instrumental action within mouse prefrontal cortex. *Eur J Neurosci*. 2010; 32:1726–1734. [PubMed: 21044173]
- Gourley SL, et al. Action control is mediated by prefrontal BDNF and glucocorticoid receptor binding. *Proc Natl Acad Sci U S A*. 2012a; 109:20714–20719. [PubMed: 23185000]
- Gourley SL, Olevska A, Warren MS, Taylor JR, Koleske AJ. Arg kinase regulates prefrontal dendritic spine refinement and cocaine-induced plasticity. *J Neurosci*. 2012b; 37:2314–2323. [PubMed: 22396406]
- Hammond LJ. The effect of contingency upon the appetitive conditioning of free-operant behavior. *J Exp Analysis Behav*. 1980; 34:297–304.
- Jetsch JD, Taylor JR. Impulsivity resulting from frontostriatal dysfunction in drug abuse: Implications for the control of behavior by reward-related stimuli. *Psychopharmacology*. 1999; 146:373–390. [PubMed: 10550488]
- Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex*. 2003; 13:400–408. [PubMed: 12631569]

- Kimch EY, Torregrossa MM, Taylor JR, Laubach M. Neuronal correlates of instrumental learning in the dorsal striatum. *J Neurophysiol.* 2009; 102:475–489. [PubMed: 19439679]
- Li C, Dabrowska J, Hazra R, Rainnie DG. Synergistic activation of dopamine D1 and TrkB receptors mediate gain control of synaptic plasticity in the basolateral amygdala. *PLoS One.* 2011; 6:e26065. [PubMed: 22022509]
- Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G. The impact of orbitofrontal dysfunction on cocaine addiction. *Nat Neurosci.* 2012; 15:358–366. [PubMed: 22267164]
- Ostlund SB, Balleine BW. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J Neurosci.* 2007; 27:4819–4825. [PubMed: 17475789]
- Paxinos, G.; Franklin, KBJ. *The mouse brain in stereotaxic coordinates.* Academic; San Diego: 2003.
- Reep RL, Corwin JV, King V. Neuronal connections of orbital cortex in rats: topography of cortical and thalamic afferents. *Exp Brain Res.* 1996; 111:215–232. [PubMed: 8891652]
- Reep RL, Cheatwood JL, Corwin JV. The associative striatum: organization of cortical projections to the dorsocentral striatum in rats. *J Comp Neurol.* 2003; 467:271–292. [PubMed: 14608594]
- Rios M, Fan G, Fekete C, Kelly J, Bates B, Kuehn R, Lechan RM, Jaenisch R. Conditional deletion of brain-derived neurotrophic factor in the postnatal brain leads to obesity and hyperactivity. *Mol Endocrinol.* 2001; 15:1748–1757. [PubMed: 11579207]
- Rushworth MF, Behrens TE, Rudebeck PH, Walton ME. Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn Sci.* 2007; 11:168–176. [PubMed: 17337237]
- Schilman EA, Uylings HB, Galis-deGraaf Y, Joel D, Groenewegen HJ. The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neurosci Lett.* 2008; 432:40–45. [PubMed: 18248891]
- Schoenbaum G, Roesch MR, Stalnaker TA, Takahashi YK. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat Rev Neurosci.* 2009; 10:885–892. [PubMed: 19904278]
- Shen K, Cowen CW. Guidance molecules in synapse formation and plasticity. *Cold Spring Harb Perspect Biol.* 2010; 2(4)
- Sul JH, Kim H, Huh N, Lee D, Jung MW. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron.* 2010; 66:449–460. [PubMed: 20471357]
- Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron.* 2009; 62:269–280. [PubMed: 19409271]
- Thoenen H. Neurotrophins and neuronal plasticity. *Science.* 1995; 270:593–598. [PubMed: 7570017]
- Thomas TM, Smith Y, Levey AI, Hirsch SM. Cortical inputs to m2-immunoreactive striatal interneurons in rat and monkey. *Synapse.* 2000; 37:252–261. [PubMed: 10891862]
- Vigers AJ, Amin DS, Talley-Farnham T, Gorski JA, Xu B, Jones KR. Sustained expression of brain-derived neurotrophic factor is required for maintenance of dendritic spine and normal behavior. *Neuroscience.* 2012; 212:1–18. [PubMed: 22542678]
- Yin HH, Knowlton BJ, Balleine BW. Lesions of the dorsolateral striatum preserve outcomes expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci.* 2004; 19:181–189. [PubMed: 14750976]
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci.* 2005; 22:513–523. [PubMed: 16045504]
- Yin HH, Knowlton BJ, Balleine BW. Inactivation of the dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav Brain Res.* 2006; 166:189–196. [PubMed: 16153716]
- Yin HH, Ostlund SB, Balleine BW. Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur J Neurosci.* 2008; 28:1437–1448. [PubMed: 18793321]
- Yin HH, Mulcare SP, Hilario MR, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, Costa RM. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci.* 2009; 12:333–341. [PubMed: 19198605]



**Figure 1. oPFC BDNF and oPFC-ventrolateral striatum interactions regulate action-outcome associative conditioning**

(a) Composites represent virus and lesion spreads in these experiments, with gray representing the largest spread and black the smallest. At bottom left, representative viral-mediated EGFP expression in the oPFC (green). Bottom right, excitotoxic lateral striatum lesion (red) counterstained with NeuN (green).

(b) oPFC-targeted *Bdnf* knockdown and oPFC-ventrolateral striatum disconnection impaired instrumental response acquisition, particularly when the reinforcement schedule escalated from a fixed ratio (FR) to random interval (RI) schedule of reinforcement in the final 2 conditioning sessions.

(c) Mice with oPFC-targeted *Bdnf* knockdown and oPFC-ventrolateral striatum disconnection were insensitive to action-outcome contingency degradation: While control groups developed a preference for the non-degraded instrumental aperture after action-outcome contingency degradation, response rates in knockdown and disconnection mice did not vary from baseline (compare to final session in b).

(d) Moreover, response acquisition in an instrumental reversal task sensitive to oPFC lesions was also impaired.

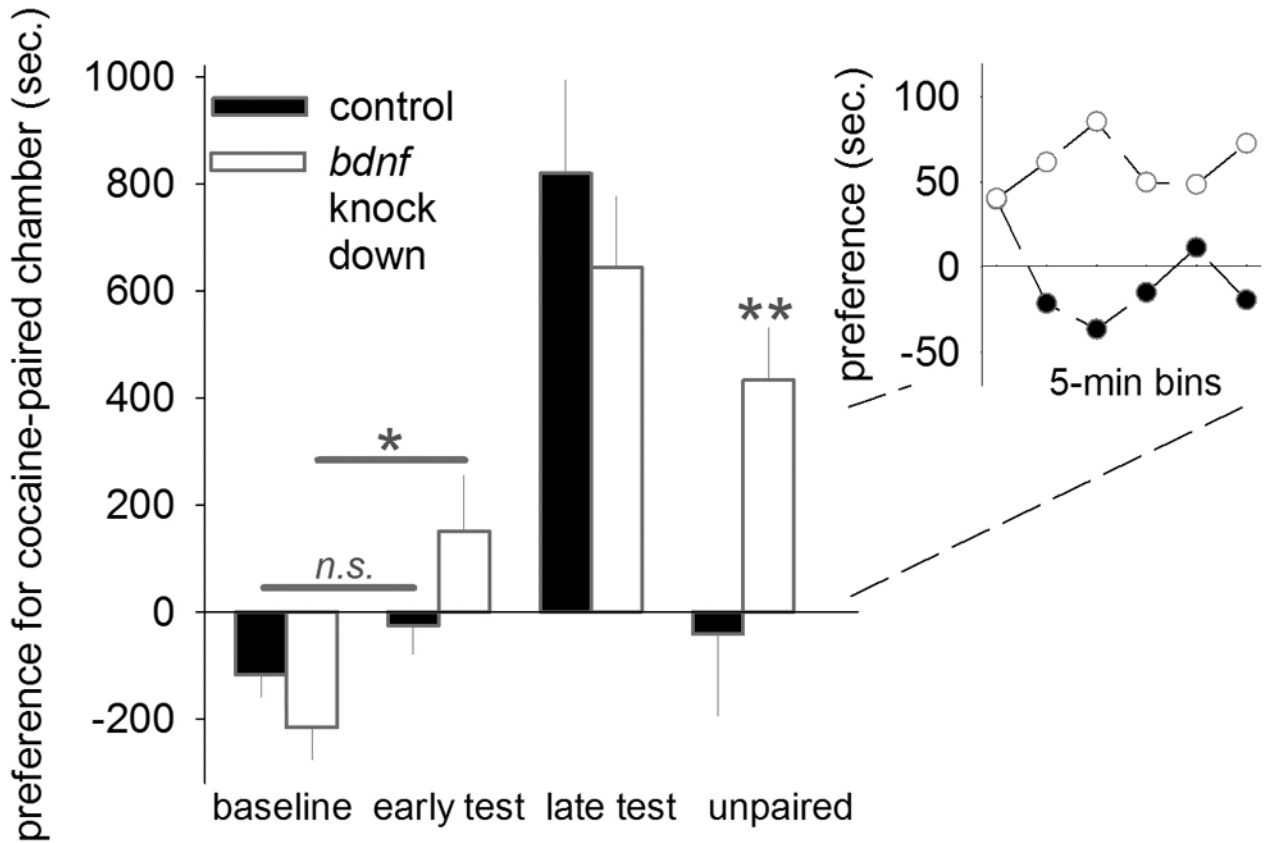
(e) These response patterns are shown again with the addition of responding on the previously reinforced aperture (“extinction”). As is summarized in the right panel, all mice initially respond preferentially on the previously reinforced aperture during the first session, but by the fourth reversal session, control and ipsilateral mice respond preferentially on the newly reinforced aperture, while the contralateral and *Bdnf* knockdown groups fail to differentiate between the reinforced and extinguished apertures.

(f) Orbitofrontal cortical *Bdnf* knockdown reduced BDNF expression in the downstream lateral striatum and amygdala, and (g) Striatal c-Fos expression predicted response rates on the reinforced aperture at the end of instrumental training (individual mice are represented; black=control group).

(h) Despite these deficiencies, both *Bdnf* knockdown and contralateral lesions spared sensitivity to satiety-specific outcome devaluation, which decreased response rates in all groups.

(i) Our model is shown: We hypothesize that while the lateral striatum coordinates stimulus-response habits as been previously reported, the ventrolateral compartment regulates, under certain circumstances, outcome-based decision-making. Means+SEMs. \* $p < 0.05$ ; \*\* $p < 0.001$ .





**Figure 2. oPFC BDNF regulates cocaine-conditioned place preference**

*Bdnf* knockdown enhanced sensitivity to cocaine-associated stimuli, as evidenced by increased place preference relative to baseline after only 2 cocaine-place pairings (“early test”). All mice ultimately acquired cocaine-conditioned place preference (“late test”), but preference failed to normalize in *Bdnf*-deficient mice when saline replaced cocaine (“unpaired”). Inset: preference in 5-min bins during the unpaired test. Means+SEMs. \* $p < 0.05$ ; \*\* $p < 0.001$ .