



## ARTICLE

# Medial orbitofrontal cortex dopamine D<sub>1</sub>/D<sub>2</sub> receptors differentially modulate distinct forms of probabilistic decision-making

Nicole L. Jenni<sup>1,2</sup>, Yi Tao Li<sup>1,2</sup> and Stan B. Floresco<sup>1,2</sup>

Efficient decision-making involves weighing the costs and benefits associated with different actions and outcomes to maximize long-term utility. The medial orbitofrontal cortex (mOFC) has been implicated in guiding choice in situations involving reward uncertainty, as inactivation in rats alters choice involving probabilistic rewards. The mOFC receives considerable dopaminergic input, yet how dopamine (DA) modulates mOFC function has been virtually unexplored. Here, we assessed how mOFC D<sub>1</sub> and D<sub>2</sub> receptors modulate two forms of reward seeking mediated by this region, probabilistic reversal learning and probabilistic discounting. Separate groups of well-trained rats received intra-mOFC microinfusions of selective D<sub>1</sub> or D<sub>2</sub> antagonists or agonists prior to task performance. mOFC D<sub>1</sub> and D<sub>2</sub> blockade had opposing effects on performance during probabilistic reversal learning and probabilistic discounting. D<sub>1</sub> blockade impaired, while D<sub>2</sub> blockade increased the number of reversals completed, both mediated by changes in errors and negative feedback sensitivity apparent during the initial discrimination of the task, which suggests changes in probabilistic reinforcement learning rather than flexibility. Similarly, D<sub>1</sub> blockade reduced, while D<sub>2</sub> blockade increased preference for larger/risky rewards. Excess D<sub>1</sub> stimulation had no effect on either task, while excessive D<sub>2</sub> stimulation impaired probabilistic reversal performance, and reduced both profitable risky choice and overall task engagement. These findings highlight a previously uncharacterized role for mOFC DA, showing that D<sub>1</sub> and D<sub>2</sub> receptors play dissociable and opposing roles in different forms of reward-related action selection. Elucidating how DA biases behavior in these situations will expand our understanding of the mechanisms regulating optimal and aberrant decision-making.

*Neuropsychopharmacology* (2021) 46:1240–1251; <https://doi.org/10.1038/s41386-020-00931-1>

## INTRODUCTION

The orbitofrontal cortex (OFC) is a key node within broader cortico-limbic-striatal circuitry that refines goal-directed behavior guided by the perceived value of action outcomes. A considerable body of evidence has shown functional specializations within different subregions of the OFC [1], with the lateral and medial OFC (mOFC) displaying distinct patterns of cortical and subcortical connectivity [2–5]. Preclinical studies of OFC function in rodents have focused primarily on the lateral portion, whereas comparatively fewer studies have explored the contribution of the mOFC to cognition and behavior [1, 6–8]. An emerging hypothesis from these studies is that the mOFC refines goal-directed action selection by retrieving representations of the estimated value of different action outcomes [8–12]. Lesions or inactivations of the mOFC can lead to adoption of behavioral strategies based more on immediate or observable reward feedback, as opposed to internal representations of outcome value shaped by reward history [7–9, 12, 13].

Choice situations involving reward uncertainty is one form of decision-making where value representations are labile, and as such, immediate reward feedback is not always the best indicator of optimal future choices [9, 14–16]. The mOFC plays a key role in guiding these types of decisions. Inactivation of this region biases

choice toward larger, uncertain rewards on a probabilistic discounting task irrespective of whether reward probabilities were initially high and decreased over the session or vice versa [9]. This effect was associated with increased win-stay behavior, in that choice was more heavily influenced by immediate reward feedback, rather than modifying choice biases based on a protracted accounting of reward history. Inactivating the mOFC also impairs probabilistic reversal learning, but either has no effect or actually improves deterministic reversal learning [7, 17], likely because choices in the latter situation can be made based on immediately observable outcomes, and would not require the retrieval of an internal representation of “correct” options.

Mesocortical dopamine (DA) has long been implicated in mediating executive functions such as selective attention, working memory, and cognitive flexibility [18, 19] and has more recently been shown to play a critical role in refining decisions involving reward uncertainty in a complex and receptor-specific manner. For example, DA acting on D<sub>1</sub> or D<sub>2</sub> receptors within the prelimbic region of the medial prefrontal cortex, acts on dissociable networks of prefrontal neurons to either bias choice toward larger, risky rewards, or facilitate flexible adjustment in choice biases [20, 21]. On the other hand, systemic blockade of D<sub>1</sub> or D<sub>2</sub> receptors did not alter performance of a probabilistic reversal task,

<sup>1</sup>Department of Psychology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada and <sup>2</sup>Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Correspondence: Stan B. Floresco (floresco@psych.ubc.ca)

Received: 3 August 2020 Revised: 10 November 2020 Accepted: 27 November 2020

Published online: 15 January 2021

but stimulation of D<sub>2</sub> receptors impaired task performance and altered negative feedback learning [22, 23]. These latter effects appeared to be driven in part by activation of striatal D<sub>2</sub> receptors [22] yet how mesocortical DA activity may influence probabilistic reinforcement learning has not been investigated.

Studies examining mesocortical DA modulation of cognition and decision-making have focused almost exclusively on terminal regions within the medial prefrontal cortex (i.e., prelimbic/infralimbic regions). However, the mOFC also receives dopaminergic input from the ventral tegmental area [24, 25] and is interconnected with key nodes within the DA decision circuitry, including the basolateral amygdala, the ventral striatum, and other prefrontal regions [2, 4]. In light of this, it is particularly surprising that there is a dearth of studies on how DA within the mOFC may modulate reward-related behaviors. Blockade of mOFC D<sub>1</sub> receptors blunted reinstatement of cocaine seeking [26] and reduced effort-related responding for food delivered on a progressive ratio [27]. Yet, how mOFC DA transmission modulates more complex forms of choice behavior entailing cost–benefit analyses related to reward uncertainty and magnitude is unknown. To address this gap in the literature, we conducted a thorough characterization of how reducing or increasing activity of mOFC DA D<sub>1</sub>-like and D<sub>2</sub>-like receptors influences two distinct forms of decision-making involving reward uncertainty. Probabilistic reversal learning assessed the contribution of these receptors to flexible action selection in response to changes in probabilistic reward contingencies, and probabilistic discounting assessed how these receptors influence risk/reward decision-making involving choice between rewards of different magnitude and probability.

## MATERIALS AND METHODS

### Animals

Adult male Long-Evans rats (Charles River Laboratories) weighing 225–300 g were food restricted for the experiment. Female rats were not included in this study. Some studies have shown that females tend to be more risk averse on probabilistic discounting tasks compared to males [28, 29]. However, despite potential baseline differences, dopaminergic manipulations such as amphetamine treatment induce similar changes in behavior in both sexes [29]. Details on housing conditions are described in Supplementary Materials and Methods. All experiments were in accordance with the Canadian Council on Animal Care guidelines regarding appropriate and ethical treatment of animals and were approved by the Animal Care Centre at the University of British Columbia.

### Initial training

Prior to training on the main task, rats received basic lever press training in operant chambers, and then retractable lever press training, before being trained on either the probabilistic reversal or discounting tasks. Details on the chambers and each step of the training procedures are described in Supplementary Methods.

**Probabilistic reversal learning.** This task was modified from procedures described by [30] and was identical to those described in previous studies [7]. At the start of each 200-trial session, one lever was designated the “correct” lever and the other “incorrect” that when selected would deliver one reward pellet on 80/20% of trials, respectively. Every 15 s, both levers were inserted, and rats had 10 s to make a choice (otherwise scored as an omission). Levers retracted after each choice or omission. Following eight consecutive “correct” completed trials (irrespective of omissions), a “reversal” was scored, reward contingencies switched, and this continued for 200 trials (Fig. 1A). Squads of rats were trained until they displayed stable performance for 3 consecutive days, after which microinfusion drug tests commenced. Additional details are described in Supplementary Materials and Methods.

**Probabilistic discounting task.** This task was modified from procedures previously described previously [9, 31]. Rats were required to choose between two levers, one designated the large/risky option, and the other the small/certain option. Choice of the small/certain option delivered one pellet with 100% probability. Choice of the large/risky option delivered four pellets with a probability that decreased systematically across five blocks of eight forced-choice trials, followed by ten free-choice trials (100, 50, 25, 12.5, 6.25%, 90 trials total; Fig. 1B). Every 35 s, one or both levers were inserted into the chamber and rats had 10 s to make a choice. Levers retracted after each choice or omission. Rats were trained until they displayed stable choice behavior after which they received surgery. They were retrained on the task prior to receiving microinfusion drug tests. Additional details are described in Supplementary Materials and Methods.

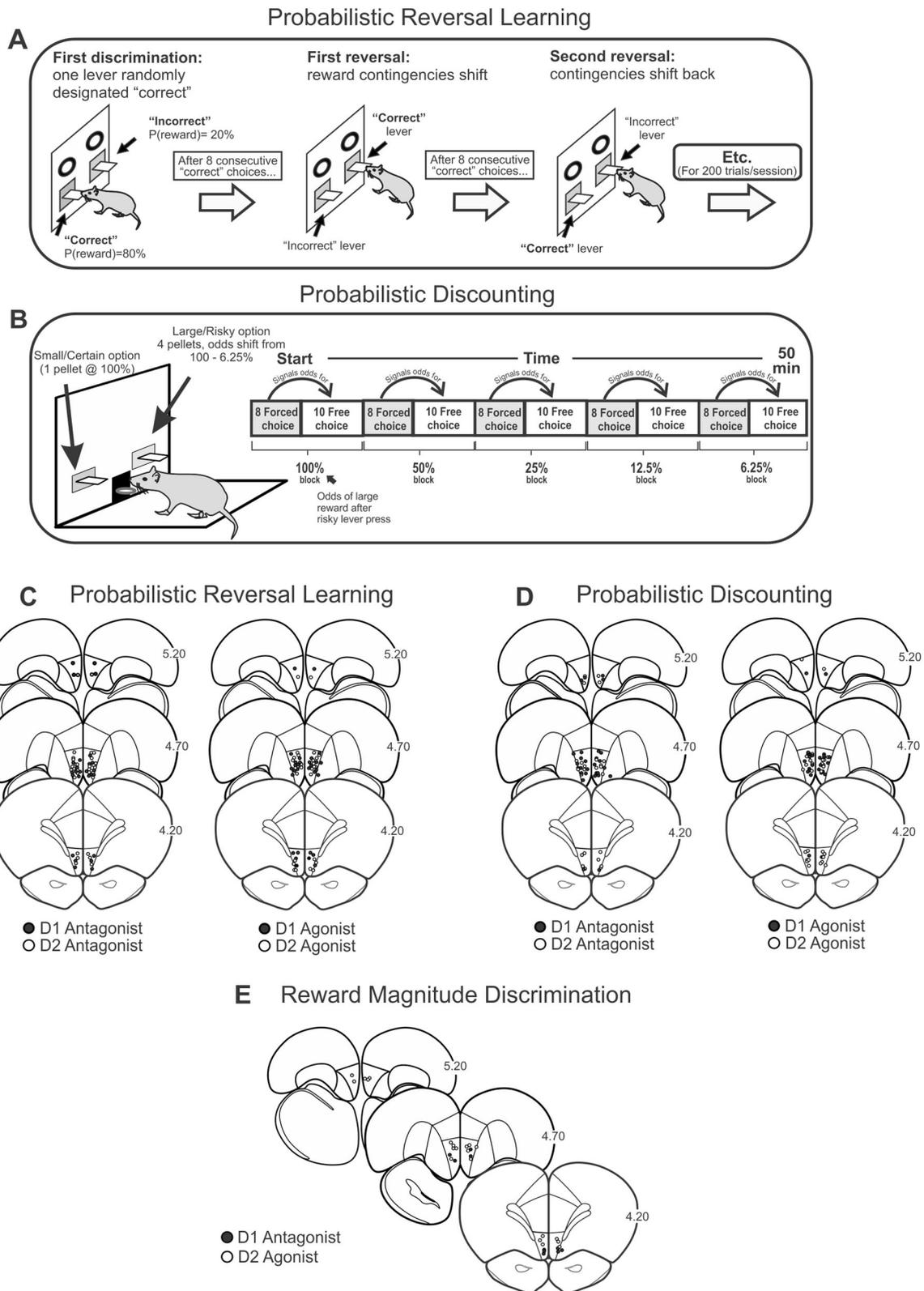
**Reward magnitude discrimination.** We determined a priori that if a manipulation reduced risky choice on the discounting task, we would test the effect of that same manipulation on a separate group of rats trained on a reward magnitude discrimination task. This task consisted of four blocks of two forced-choice trials followed by ten free-choice trials where choice of the large reward lever delivered four pellets, and choice of the small reward lever delivered one pellet, both with 100% probability. Additional details are described in Supplementary Materials and Methods.

**Surgery.** Rats were implanted with bilateral 23-gauge guide cannulae targeted above the mOFC [AP = +4.5 mm; ML = ±0.7 mm from bregma; DV = –3.3 mm from dura] following standard stereotaxic procedures, which have been described by our group previously [9, 31]. Rats were given at least 1 week to recover. Additional details are described in Supplementary Materials and Methods.

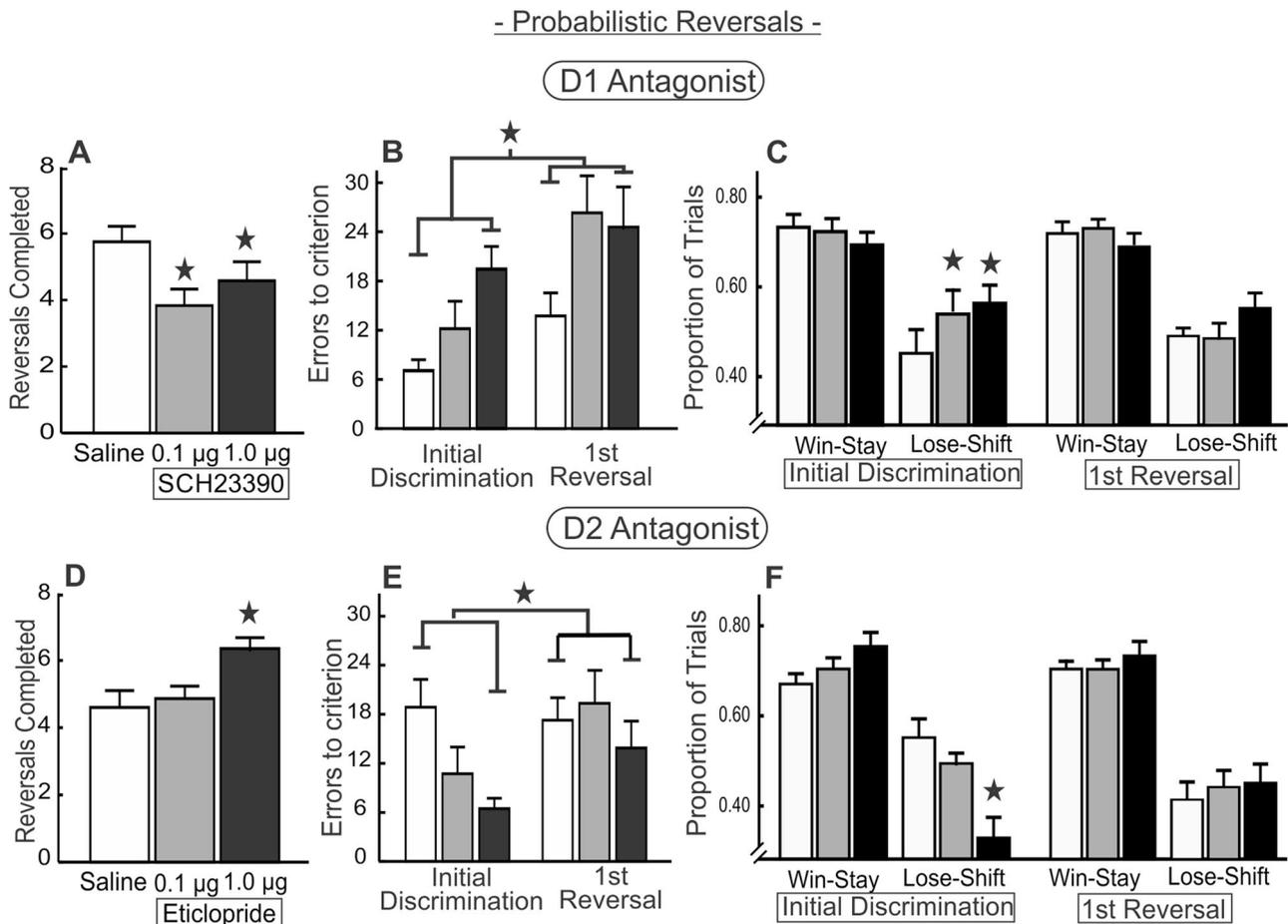
**Drugs and microinfusion procedure.** Once stable choice behavior was established, rats in each experiment were assigned to a drug group and received their first of three counterbalanced intra-mOFC microinfusions of either vehicle, the low or high dose of the drug, administered in 0.5 µl, 10 min prior to testing. Following the first drug tests, rats received at least two drug-free retraining sessions before the next test. The drugs (all obtained from Sigma-Aldrich) and doses used were: D<sub>1</sub> antagonist SCH 23390 (0.1, 1.0 µg); D<sub>2</sub> antagonist eticlopride (0.1, 1.0 µg); D<sub>1</sub> agonist SKF 81297 (0.1, 0.4 µg); and D<sub>2</sub> agonist quinpirole (1, 10 µg), all dissolved in 0.9% saline. These doses were chosen based on previous studies that showed them to be effective at altering probabilistic discounting and other forms of executive functioning when infused bilaterally into the medial PFC [18, 32, 33]. Additional details are described in Supplementary Materials and Methods.

**Histology.** Rats were euthanized in a CO<sub>2</sub> chamber. Brains were fixed in 4% formalin. Each brain was frozen, sliced in 50 µm sections, and Nissl stained with cresyl violet. Placements are displayed in Fig. 1C–E. Rats whose placements resided outside the defined borders of the mOFC, and that encroached on the prelimbic cortex or penetrated into the ventral fissure were removed from the analysis. Across all groups, 125 rats are included in the analyses and 38 rats were excluded due to their infusions being too dorsal or ventral to the defined borders of the mOFC.

**Data analysis.** **Probabilistic reversal learning.** The primary outcome variable was the number of reversals completed, with secondary analyses conducted to clarify how a particular treatment affected win–stay and lose–shift ratios, errors during the initial discrimination and first reversal, response latencies, trial omissions and locomotion. Most variables were analyzed with one-way repeated measures ANOVAs. Analysis of the errors involved two-way ANOVAs that included treatment and task



**Fig. 1 Task procedures and histology.** **A** Probabilistic reversal learning task. At the start of a session, rats select one of two levers randomly designated as “correct” or “incorrect.” Each correct/incorrect choice is rewarded 80/20% of the time, respectively (left). After eight consecutive correct choices, the reward contingencies are reversed (middle) and this occurs again after another 8 consecutive correction choices for 200 trials (right). **B** Cost/benefit contingencies associated with responding on either lever on the probabilistic discounting task. Right, format of the sequence of forced and free-choice trials within each probability block in which the odds of obtaining the larger reward decreased from 100 to 6.25% across five blocks of trials. Schematic of coronal sections of the rat brain showing location of acceptable infusions in the mOFC for rats in the **C** probabilistic reversal learning, **D** probabilistic discounting, and **E** reward magnitude discrimination experiments.



**Fig. 2 Blockade of D<sub>1</sub> vs D<sub>2</sub> receptors within mOFC differentially alters probabilistic reversal learning.** **A** Infusion of SCH 23390 into the mOFC ( $n = 13$ ) reduced the number of reversals completed. **B** Errors to achieve criterion performance during the initial discrimination and first reversal phases. D<sub>1</sub> blockade increased errors made during the initial discrimination and following the first reversal. **C** D<sub>1</sub> blockade increased lose–shift behavior during the initial discrimination but did not affect win–stay or lose–shift ratios following the first reversal. **D** Infusion of eticlopride into the mOFC ( $n = 15$ ) increased the number of reversals completed. **E** D<sub>2</sub> blockade reduced errors during the initial discrimination and first reversal phases. **F** D<sub>2</sub> blockade reduced lose–shift behavior during the initial discrimination of the task. For this and all other figures, error bars are SEM, stars denotes  $p < 0.05$  compared to saline and brackets above the bars in **B** and **E** indicate significant main effect of treatment independent of phase.

phase (initial discrimination, first reversal) as within-subjects factors. Win–stay/lose–shift ratios were analyzed with an ANOVA model that included treatment, task phase, response type (win–stay/lose–shift), and choice type (correct vs incorrect trials) as factors (see Supplementary Materials and Methods). However, there were no interactions with the choice type factor across any treatments (all  $F_s \leq 2.02$   $p_s > 0.15$ ). Accordingly, for all the analyses presented, win–stay/lose–shift ratios represent these types of responses collapsed across correct and incorrect choices. Note that this approach only examined how the most recent outcome of a choice affected subsequent choice, and may only serve as partial indicators of learning rates or reward/negative feedback sensitivity given the probabilistic nature of the tasks [34].

**Probabilistic discounting and reward magnitude discrimination.** The primary outcome variable was the percentage of choices of the large/risky option during each block of trials, analyzed using a two-way ANOVA (treatment  $\times$  trial block as two within-subjects factors). Significant effect on risky choice triggers supplementary analyses on how these treatments altered win–stay or lose–shift behavior to ascertain how rewarded/nonrewarded risky choices influence subsequent action selection.

For both tasks, response latencies, trial omissions, and locomotion (photobeam breaks) were analyzed with one-way ANOVAs and all follow-up multiple comparisons were made using

Dunnett’s test. Additional details concerning variable calculations and data analysis are presented in Supplementary Materials and Methods.

## RESULTS

### mOFC DA modulation of probabilistic reversal learning

Previous studies have shown that inactivation of the mOFC impaired probabilistic reversal performance [7, 34]. Here, we determined how reducing or increasing activity of D<sub>1</sub> or D<sub>2</sub> receptors in the mOFC alters this form of reward-related decision-making.

**D<sub>1</sub> receptor blockade.** In 13 well-trained rats with acceptable placements (Fig. 1C, left), infusions of both the low and high dose of SCH 23390 impaired probabilistic reversal performance, indexed by a reduction in the number of reversals completed ( $F(2,24) = 4.37$ ,  $p = 0.02$  and Dunnett’s,  $p < 0.05$ , Fig. 2A). Additional analyses probed whether this reflected an impairment during the reversal shift, or a more fundamental impairment in probabilistic reinforcement learning, by analyzing the errors made during the initial discrimination and following the first reversal. This yielded a main effect of treatment ( $F(2,24) = 4.97$ ,  $p = 0.02$ ), and phase ( $F(1,12) = 13.62$ ,  $p = 0.003$ ), but no treatment  $\times$  phase interaction

**Table 1.** Win–stay and lose–shift ratios displayed following D<sub>1</sub> or D<sub>2</sub> blockade and stimulation during probabilistic reversal learning, computed across the entire session, and averaged across correct and incorrect choices.

	Saline		Low dose		High dose	
	Win–stay	Lose–shift	Win–stay	Lose–shift	Win–stay	Lose–shift
D1 antagonist (SCH 23390)	0.72 ± 0.02	0.47 ± 0.02	0.70 ± 0.02	0.55 ± 0.02	0.69 ± 0.03	0.50 ± 0.03
D2 antagonist (eticlopride)	0.71 ± 0.01	0.46 ± 0.02	0.72 ± 0.01	0.47 ± 0.01	0.73 ± 0.02	0.45 ± 0.02
D1 agonist (SKF 81297)	0.74 ± 0.01	0.49 ± 0.02	0.79 ± 0.04	0.50 ± 0.05	0.76 ± 0.01	0.48 ± 0.02
D2 agonist (quinpirole)	0.76 ± 0.02	0.48 ± 0.02	0.75 ± 0.02	0.50 ± 0.03	0.72 ± 0.02	0.45 ± 0.03

Ratios are displayed for correct and incorrect trials across the whole session. Values displayed are mean ± SEM.

( $F(2,24) = 0.80, p = 0.46$ ; Fig. 2B), precluding us from comparing errors across treatments during each individual phase. Multiple comparisons revealed that only the high dose of SCH 23390 increased errors during these phases of the task ( $p < 0.05$ ), indicating that this dose impaired performance during the initial discrimination and that this persisted through the first reversal.

Analysis of win–stay/lose–shift data over the entire session yielded no main effect of treatment ( $F(2,24) = 1.37, p = 0.27$ ) or treatment × response type interaction ( $F(2,24) = 2.52, p = 0.10$ ; Table 1). However, given that the impairment in performance was driven by increased errors during the initial discrimination and first reversal phases, we conducted a more targeted analysis of win–stay/lose–shift behavior during these task phases. We analyzed these ratios with a three-way ANOVA with treatment, response (win–stay, lose–shift), and task phase (initial discrimination, first reversal) as three within-subjects factors. This analysis yielded a three-way treatment × response type × task phase interaction ( $F(2,24) = 5.14, p = 0.01$ ). Partitioning of this three-way interaction further revealed that during the initial discrimination, both doses of SCH 23390 increased lose–shift behavior ( $F(2,24) = 4.99, p = 0.01$  and Dunnett’s,  $p < 0.05$ , Fig. 2C, left) without affecting win–stay behavior ( $F(2,24) = 0.91, p = 0.42$ ). In comparison, these treatments did not affect win–stay/lose–shift behavior during the first reversal (main effect  $F(2,24) = 0.39, p = 0.68$ , treatment × phase  $F(2,24) = 0.13, p = 0.88$ ; Fig. 2C, right). Blockade of mOFC D<sub>1</sub> receptors had no effect on omissions, locomotion, or response latencies (all  $F_s < 1.72$ , all  $p_s > 0.20$ ; Table 2). Collectively, these data suggest that mOFC D<sub>1</sub> activity facilitates the use of probabilistic reward feedback to discriminate between actions that are more vs less profitable over time. Furthermore, the observation that D<sub>1</sub> blockade impaired performance from the initial discrimination suggests these impairments likely reflect a diminished capacity to maintain choice biases toward more profitable options.

**D<sub>2</sub> receptor blockade.** Fifteen rats received infusions of eticlopride localized within the mOFC (Fig. 1C, left). In contrast to the effects of D<sub>1</sub> blockade, treatment with the high dose of a D<sub>2</sub> antagonist actually increased the number of reversals completed ( $F(2,28) = 5.38, p = 0.01$  and Dunnett’s,  $p < 0.05$ ; Fig. 2D). Analysis of errors revealed a main effect of treatment ( $F(2,28) = 4.25, p = 0.02$ ) but no treatment × phase interaction ( $F(2,28) = 1.83, p = 0.18$ ; Fig. 2E), with rats committing fewer errors after treatment with the high dose ( $p < 0.05$ ). Despite the lack of interaction, visual inspection of the data shows this effect was driven primarily by a reduction in errors committed during the initial discrimination, suggesting the increase in reversals may be mediated by a more general refinement in probabilistic reinforcement learning rather than improved flexibility.

Once again, there was no effect of treatment on win–stay/lose–shift ratios computed over the entire session (main effect  $F(2,28) = 0.36, p = 0.70$ ; treatment × response type  $F(2,28) = 1.38, p = 0.27$ ; Table 1). However, a targeted analysis on the first two task

phases yielded a treatment × response type × phase interaction ( $F(2,28) = 11.61, p < 0.001$ ). This was driven by a reduction in lose–shift behavior during the initial discrimination following treatment with the high dose of eticlopride ( $F(2,28) = 9.25, p = 0.001$  and Dunnett’s,  $p < 0.05$ ), combined with a trend toward an increase in win–stay behavior during this phase ( $F(2,28) = 3.13, p = 0.06$ ). Conversely, during the first reversal, these measures were unaltered by mOFC D<sub>2</sub> blockade (main effect treatment:  $F(2,28) = 0.70, p = 0.51$ ; treatment × response type interaction:  $F(2,28) = 0.09, p = 0.91$ ; Fig. 2F). D<sub>2</sub> blockade had no effect on response latency, trial omissions, or locomotion (all  $F_s < 0.82$ , all  $p_s > 0.45$ ; Table 2). Collectively, these data show that D<sub>2</sub> blockade facilitates reversal performance, by reducing errors and the tendency to shift choice after a nonrewarded action during the initial discrimination of the task. In contrast to mOFC D<sub>1</sub> receptors that appear to facilitate maintenance of choice of more profitable choices, mOFC D<sub>2</sub> receptor function may mitigate this bias in favor of exploring alternative options in response to nonrewarded actions.

When comparing the effects of SCH 23390 and eticlopride, we noticed that rats in the D<sub>1</sub> antagonist group completed more reversals under control conditions compared to the D<sub>2</sub> antagonist group. Although this difference was not statistically significant ( $t_{(26)} = 1.66, p = 0.11$ ), we wanted to confirm that these opposing effects of D<sub>1</sub> vs D<sub>2</sub> antagonism were not artifacts attributable to baseline differences in performance. In so doing, we analyzed data from subsets of rats that were matched for control performance across the two drug groups. In this analysis, we removed the three poorest performers in the eticlopride group (completing 1–3 reversals after saline infusions), and the best performer in the D<sub>1</sub> antagonist group (eight reversals after saline). This left 12 rats in each group that displayed comparable control performance (D<sub>1</sub> group mean =  $5.50 ± 0.42$ ; D<sub>2</sub> group mean =  $5.25 ± 0.38$ ;  $t_{(22)} = 0.42, p = 0.68$ , Supplementary Fig. 1). Importantly, this subset of rats showed similar profiles of change following D<sub>1</sub> or D<sub>2</sub> antagonism in the mOFC, with a reduction in reversals following D<sub>1</sub> blockade ( $F(2,22) = 3.30, p = 0.056$ ) and an increase in reversals following D<sub>2</sub> blockade ( $F(2,22) = 3.02, p = 0.03$ , Supplementary Fig. 1A). Furthermore, there were similar effects on errors to criterion and first discrimination win–stay/lose–shift across both groups after matching for reversals completed (see Supplementary Fig. 1). Thus, the opposing effects of D<sub>1</sub> vs D<sub>2</sub> blockade on probabilistic reversal performance are unlikely attributable to differences in baseline performance across the two groups.

**D<sub>1</sub> receptor stimulation.** This experiment yielded 15 animals with accurate placements (Fig. 1C, right). Increased stimulation of mOFC D<sub>1</sub> receptors did not affect the number of reversals completed ( $F(2,28) = 5.42, p = 0.14$ ; Fig. 3A), or errors committed during the initial discrimination and first reversal phases (main effect:  $F(2,28) = 2.12, p = 0.12$ ; treatment × phase:  $F(2,28) = 0.35, p = 0.71$ ; Fig. 3B). Similarly, these treatments did not affect win–stay/lose–shift behavior over the entire session (all  $F_s < 2.63$ ,

**Table 2.** Performance measures following D<sub>1</sub> and D<sub>2</sub> blockade and stimulation during probabilistic reversal learning, probabilistic discounting, and reward magnitude discrimination.

	Saline	Low dose	High dose
<b>Probabilistic reversal</b>			
D1 antagonist (SCH 23390)			
Response latency	0.79 ± 0.1	0.73 ± 0.1	0.73 ± 0.1
Trial omissions	3.1 ± 2	1.8 ± 1	3.3 ± 2
Locomotion	1398 ± 168	1421 ± 116	1210 ± 175
D2 antagonist (eticlopride)			
Response latency	0.67 ± 0.1	0.73 ± 0.1	0.72 ± 0.1
Trial omissions	4.2 ± 3	2.4 ± 2	5.3 ± 4
Locomotion	1634 ± 169	1711 ± 124	1659 ± 211
D1 agonist (SKF 81297)			
Response latency	0.70 ± 0.1	0.60 ± 0.1	0.61 ± 0.1
Trial omissions	7.8 ± 6	1.9 ± 1	1.3 ± 1
Locomotion	1594 ± 191	1434 ± 239	1556 ± 202
D2 agonist (quinpirole)			
Response latency	0.84 ± 0.1	1.04 ± 0.2	1.34 ± 0.3
Trial omissions	9.4 ± 5	6.0 ± 2	34.8 ± 18*
Locomotion	1423 ± 191	1179 ± 184	1228 ± 159
<b>Probabilistic discounting</b>			
D1 antagonist (SCH 23390)			
Response latency	0.71 ± 0.1	0.67 ± 0.1	0.92 ± 0.1*
Trial omissions	3.8 ± 2	2.7 ± 1	6.2 ± 3
Locomotion	1698 ± 241	1531 ± 251	1453 ± 241
D2 antagonist (eticlopride)			
Response latency	0.81 ± 0.1	0.75 ± 0.1	0.95 ± 0.2
Trial omissions	5.5 ± 3	2.1 ± 1	4.2 ± 2
Locomotion	1409 ± 164	1454 ± 150	1374 ± 158
D1 agonist (SKF 81297)			
Response latency	1.08 ± 0.1	1.14 ± 0.1	1.02 ± 0.1
Trial omissions	3.6 ± 2	4.9 ± 2	3.8 ± 1
Locomotion	1441 ± 202	1101 ± 108	1399 ± 141
D2 agonist (quinpirole)			
Response latency	0.76 ± 0.1	0.91 ± 0.1	0.95 ± 0.1
Trial omissions	1.8 ± 1	1.4 ± 1	4.1 ± 2*
Locomotion	1482 ± 221	1332 ± 117	1034 ± 107*
<b>Reward magnitude</b>			
D1 antagonist (SCH 23390)			
Response latency	0.79 ± 0.1	0.75 ± 0.1	0.77 ± 0.1
Trial omissions	0 ± 0	0 ± 0	0 ± 0
Locomotion	990 ± 112	1081 ± 107	1135 ± 92
D2 agonist (quinpirole)			
Response latency	1.41 ± 0.3	1.77 ± 0.3	2.70 ± 0.5*
Trial omissions	1.6 ± 1	6.5 ± 3	17.8 ± 4*
Locomotion	678 ± 131	554 ± 109	674 ± 331

Latencies are measured in s, and locomotion is indexed by photobeam breaks. Values displayed are mean ± SEM.  
 \**p* < 0.05 compared to saline.

*D<sub>2</sub> receptor stimulation.* Fourteen rats were included in this analysis (Fig. 1C, right), revealing that the high dose of quinpirole reduced the number of reversals completed ( $F(2,26) = 5.25, p = 0.01$ ; Dunnett's  $p < 0.05$ , Fig. 3D). Notably, quinpirole also increased the omission rate at the high dose ( $F(2,26) = 3.33, p = 0.05$  and Dunnett's  $p < 0.05$ ; Table 2). To determine whether the reduction in reversals was driven by a reduction in trials completed, a supplementary analysis compared the number of reversals per 100 completed trials, computed using the formula (# of reversals/# of completed trials × 100) as we have done previously [7, 14]. When analyzing the data in this manner, the results of the overall ANOVA only approached significance (mean ± SEM; saline = 3.3 ± 0.2; low dose = 3.3 ± 0.3; high dose = 2.5 ± 0.4;  $F(2,26) = 2.79, p = 0.08$ ). However a direct comparison of these values obtained after saline vs the high dose confirmed that 10 µg quinpirole significantly reduced the number of reversals completed/100 trials relative to control treatments ( $t(13) = 2.18, p = 0.047$ ).

D<sub>2</sub> receptor stimulation had no reliable effect on errors committed during the first two task phases (main effect  $F(2,26) = 0.23, p = 0.80$ ; treatment × phase  $F(2,26) = 0.41, p = 0.67$ ; Fig. 3E). There was no effect of treatment on win–stay/lose–shift ratios computed over the entire session (main effect  $F(2,26) = 2.15, p = 0.14$ , treatment × response type ( $F(2,26) = 0.40, p = 0.67$ ), Table 1). However, the targeted analysis on the first two task phases revealed a treatment × response type × phase interaction ( $F(2,26) = 7.39, p = 0.003$ ). This was quantified by an increase in lose–shift during the initial discrimination by the low dose of quinpirole ( $F(2,26) = 5.94, p = 0.008$ , Dunnett's  $p < 0.05$ , Fig. 3F), but curiously, not the high dose, even though this dose reduced reversals completed. The lack of effect of the high dose on these measures may be related to its increase in trial omissions, which would lengthen intervals between choices and potentially diminish the influence that an outcome has on a subsequent choice. No changes in win–stay/lose–shift values were observed during the first reversal ( $F_s < 1.71, p_s > 0.20$ , Fig. 3F). In addition to the increase in omissions mentioned above, mOFC D<sub>2</sub> stimulation tended to reduce locomotion and increase response latencies although the overall ANOVA of these data did not achieve statistical significance (locomotion:  $F(2,26) = 3.04, p = 0.065$ ; latencies:  $F(2,26) = 3.09, p = 0.062$ , Table 2). Thus, excessive mOFC D<sub>2</sub> activation with a high dose of quinpirole impaired both task engagement and reversal efficiency. Taken together with the D<sub>2</sub> antagonist data, it appears mOFC D<sub>2</sub> receptors promote shifts in choice behavior in response to losses.

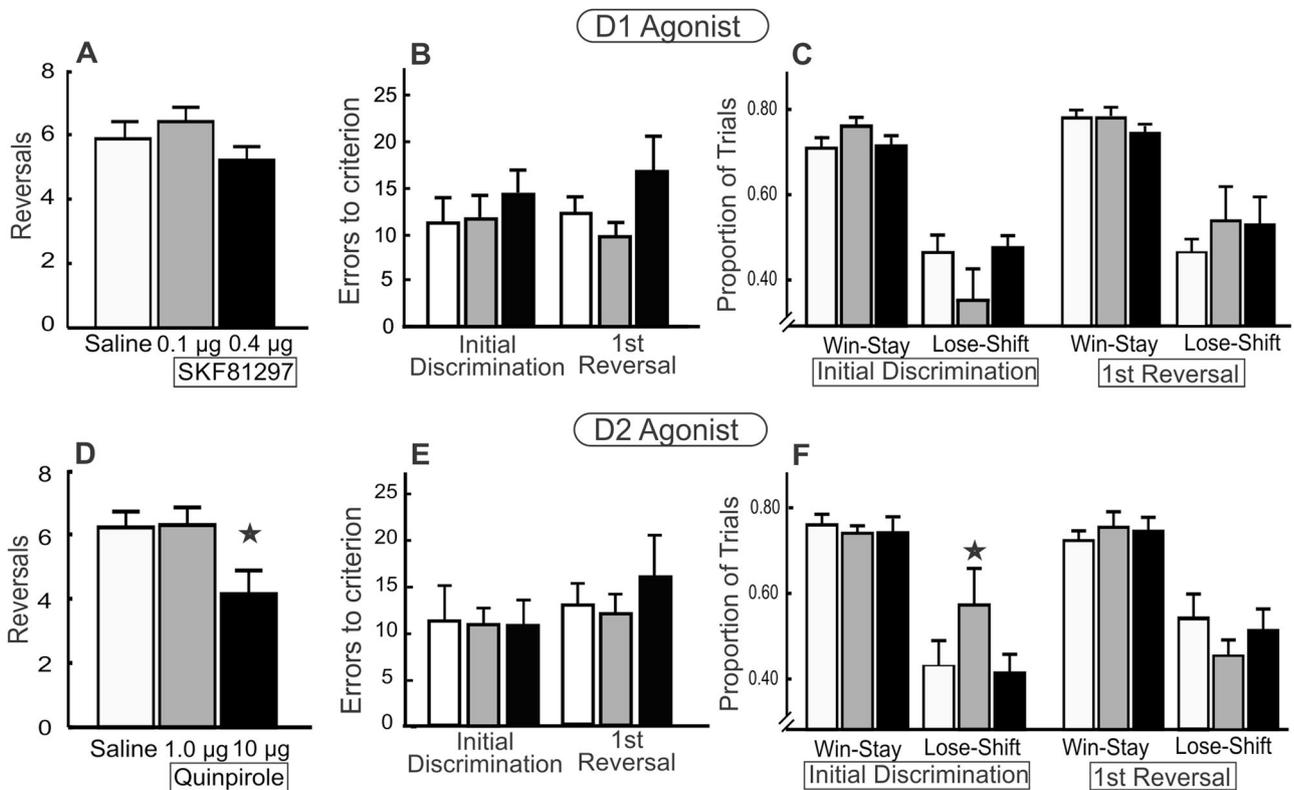
**mOFC DA modulation of probabilistic discounting**

Inactivation of the mOFC increased choice of large, uncertain rewards vs smaller certain ones on a probabilistic discounting task [9]. This experiment sought to determine how D<sub>1</sub> or D<sub>2</sub> receptor activity within the mOFC may influence this form of decision-making.

*D<sub>1</sub> receptor blockade.* Fourteen rats with acceptable placements were included in the analysis (Fig. 1D, left). Analysis of the choice data revealed that mOFC D<sub>1</sub> blockade reduced preference for larger/risky rewards, as indicated by a main effect of treatment ( $F(2,26) = 4.56, p = 0.02$  Fig. 4A) following treatment with both low and high doses of SCH 23390 (Dunnett's  $p < 0.05$ ). The treatment × block interaction did not obtain statistical significance ( $F(8,104) = 1.80, p = 0.09$ ). Subsequent analyses probed how this reduction in risky choice may have been driven by alterations in how the outcome of a risky choice influenced the next choice. A two-way repeated measures ANOVA on win–stay/lose–shift ratios revealed a main effect of treatment ( $F(2,24) = 7.49, p = 0.003$ ), but no treatment × response type interaction ( $F(2,24) = 1.79, p = 0.19$ ; Fig. 4B). mOFC D<sub>1</sub> blockade caused an overall increase in sensitivity to both rewarded and nonrewarded choices following

all  $p_s > 0.09$ ; Table 1) or when comparing these values during the first two task phases (all  $F_s < 1.47, p_s > 0.25$ , Fig. 3C). Similarly, D<sub>1</sub> receptor stimulation had no effects on any other performance measures (all  $F_s < 2.29$ , all  $p_s > 0.12$ , Table 2).

- Probabilistic Reversals -



**Fig. 3** Effects of stimulation of mOFC D<sub>1</sub> and D<sub>2</sub> receptors on probabilistic reversal performance. **A–C** Infusion of SKF 81297 into mOFC ( $n = 15$ ) did not affect probabilistic reversal performance, errors during the initial discrimination/first reversal, nor did it influence win-stay or lose-shift behavior during the first two task phases. **D** Infusion of quinpirole ( $n = 14$ ) reduced the number of reversals completed **E** Excessive D<sub>2</sub> stimulation did not affect initial discrimination or first reversal errors. **F** The low dose of quinpirole increases lose-shift behavior during the initial discrimination.

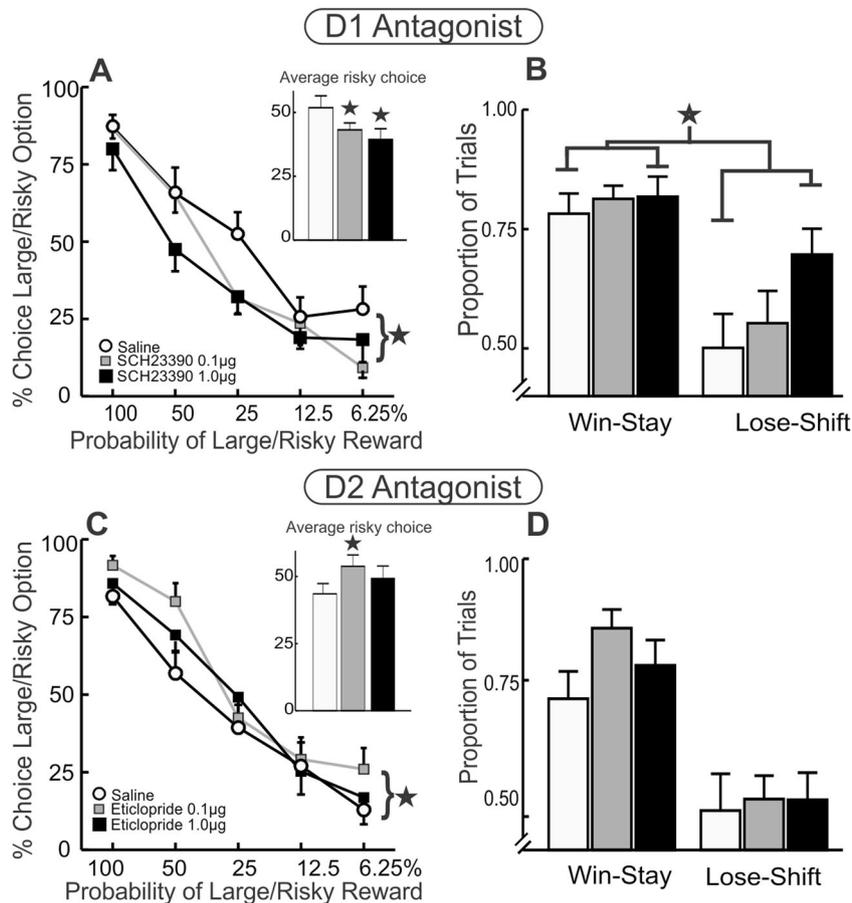
treatment with the high dose ( $p < 0.05$ ), although visual inspection of Fig. 4B suggests that the reduction in risky choice was driven primarily by increases in lose-shift behavior. D<sub>1</sub> blockade increased choice latency following the high dose ( $F(2,26) = 5.56$ ,  $p = 0.01$ ; and Dunnett's,  $p < 0.05$ ) but did not affect locomotion or trial omissions ( $F_s < 2.1$ ,  $p_s > 0.15$ ; Table 2). Collectively, these data indicate that D<sub>1</sub> receptor activity in the mOFC promotes choice of larger/risky rewards, in part by mitigating sensitivity to reward omissions. Notably, this effect is opposite to that induced by inactivation of the mOFC [9].

**D<sub>2</sub> receptor blockade.** Twelve rats with acceptable placements in the mOFC were included in the analysis (Fig. 1D, left). In contrast to the effects of D<sub>1</sub> blockade, D<sub>2</sub> antagonism increased risky choice, quantified by a main effect of treatment ( $F(2,22) = 3.82$ ,  $p = 0.04$ ; Fig. 4C), but no treatment  $\times$  block interaction ( $F(8,88) = 0.87$ ,  $p = 0.55$ ). Multiple comparisons with Dunnett's test further revealed that the 0.1  $\mu$ g dose of eticlopride induced a statistically significant increase risky choice relative to saline ( $p < 0.05$ ), whereas the effects of the higher 1.0  $\mu$ g dose did not achieve significance. Win-stay ratios were higher following D<sub>2</sub> blockade relative to control treatments, yet the overall analysis of these data failed to yield a statistically significant effect (treatment main effect  $F(2,22) = 1.59$ ,  $p = 0.23$ ; interaction  $F(2,22) = 0.95$ ,  $p = 0.40$ ; Fig. 4D). There were no effects of D<sub>2</sub> blockade on locomotion, omissions, or decision latencies (all  $F_s < 2.11$ , all  $p_s > 0.15$ ; see Table 2). Thus, similar to the effects on probabilistic reversal learning, D<sub>1</sub> vs D<sub>2</sub> receptor antagonism induced opposing effects on risk/reward decision-making.

**D<sub>1</sub> receptor stimulation.** Intra-mOFC infusions of SKF 81297 ( $n = 13$ ; Fig. 1D, right) did not affect choice behavior (main effect  $F(2,24) = 0.24$ ,  $p = 0.79$ , interaction  $F(8,96) = 1.65$ ,  $p = 0.12$ ; Fig. 5A) and thus we did not analyze how these treatments affected win-stay/lose-shift behavior. Similarly, D<sub>1</sub> stimulation had no effect on any of the other performance measures (all  $F_s < 2.68$ , all  $p_s > 0.09$ ; Table 2).

**D<sub>2</sub> receptor stimulation.** This experiment yielded 14 rats with acceptable placements (Fig. 1D, right). Analysis of choice behavior revealed a main effect of treatment ( $F(2,26) = 3.34$ ,  $p = 0.05$ ) and a treatment  $\times$  block interaction ( $F(8,104) = 2.12$ ,  $p = 0.04$ ; Fig. 5B). This interaction was driven by a reduction in risky choice during the 50% block following treatment with both doses ( $F(2,104) = 9.48$ ,  $p < 0.001$  and Dunnett's  $p < 0.05$ ), which, incidentally, was the block associated with the most uncertainty as to whether a risky choice would be rewarded or not. Choice during the other blocks did not differ significantly across treatments, including the 100% block, indicating that rats were still able to discriminate between larger vs smaller rewards (all  $F_s < 2.18$ ,  $p_s > 0.11$ ). Analysis of win-stay/lose-shift ratios further revealed a significant treatment  $\times$  response type interaction ( $F(2,26) = 9.93$ ,  $p = 0.001$ ; Fig. 5C). This was driven by both a reduction in sensitivity to rewarded risky choices (win-stay) at the high dose ( $F(2,26) = 3.39$ ,  $p = 0.05$ , Dunnett's  $p < 0.05$ ) and an increased sensitivity to nonrewarded risky choices (lose-shift) after treatment with both doses ( $F(2,26) = 8.33$ ,  $p = 0.001$  and Dunnett's,  $p < 0.05$ ). The high dose of quinpirole also increased choice latencies ( $F(2,26) = 5.63$ ,  $p = 0.009$ ), reduced locomotion ( $F(2,26) = 13.69$ ,  $p < 0.001$ ), and

- Probabilistic Discounting -



**Fig. 4 Blockade of D<sub>1</sub> and D<sub>2</sub> receptors within the mOFC differentially impairs probabilistic discounting.** **A** Percentage choice of the large/risky option following infusion of SCH 23390 into the mOFC ( $n = 14$ ) across five blocks of free-choice trials. Inset displays the main effect in the analysis, averaging total percentage of large/risky choice across the all trial blocks. Blockade of D<sub>1</sub> receptors reduced choice for the larger reward. **B** Win-stay/lose-shift ratios. D<sub>1</sub> blockade at the high dose increased sensitivity to losses and to a lesser degree, wins. **C** Percentage choice of the large/risky option following infusion of eticlopride into the mOFC ( $n = 12$ ) across five blocks of free-choice trials. Inset displays large/risky choice averaged across the all block of trials demonstrating the main effect of treatment and increase in risky choice. **D** Win-stay/lose-shift ratios following intra-mOFC eticlopride.

increased omissions ( $F(2,26) = 5.62, p = 0.009$ ) ( $p < 0.05$ , Table 2). Thus, excess activation of mOFC D<sub>2</sub> receptors again reduced task engagement, while also altering the profile of probabilistic choice. Specifically, excessive D<sub>2</sub> receptor activity attenuates bias for larger rewards when their probabilistic schedules are highly uncertain, by both dampening the impact a large reward has over subsequent choice, and increasing their sensitivity to nonrewarded actions.

**Reward magnitude discrimination**

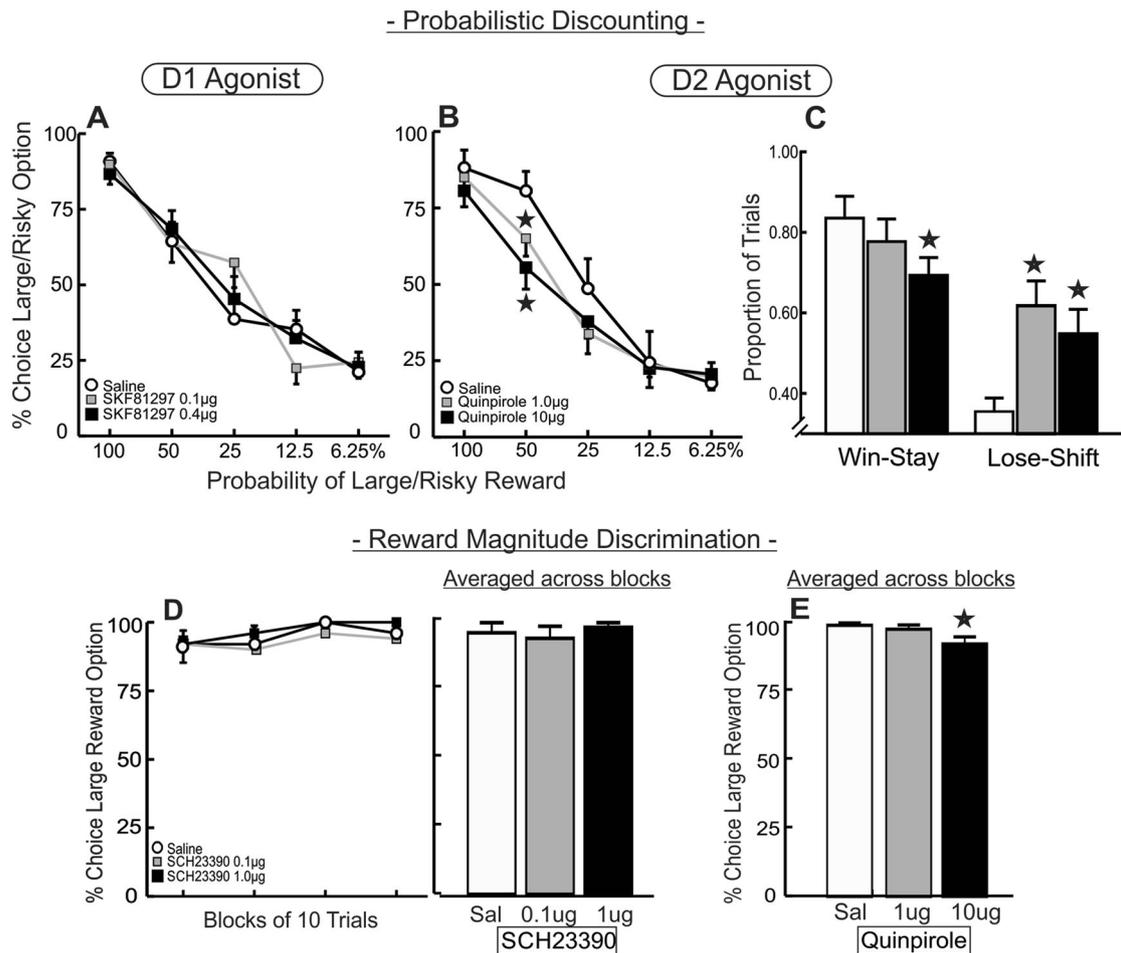
Both blockade of D<sub>1</sub> and stimulation of D<sub>2</sub> receptors in the mOFC reduced choice of larger, uncertain reward. In light of these effects, we conducted additional experiments to assess how these manipulations affect choice between a one-pellet vs four-pellet reward, both delivered with 100% certainty. The experiment entailing infusions of the D<sub>1</sub> antagonist SCH 23390 yielded five rats with accurate placements confined within the mOFC (Fig. 1E). As is clear from fig. 5D, D<sub>1</sub> blockade had no effect on choice (all  $F_s < 1.0$ , all  $p_s > 0.80$ ), and also did not affect any other performance measure (all  $F_s < 0.92$ , all  $p_s > 0.43$ ; Table 2).

For rats that received intra-mOFC infusions of quinpirole ( $n = 10$ ), we observed that these treatments yet again increased omissions ( $F(2,18) = 13.53, p < 0.001$ ; Table 2) at both doses ( $p < 0.05$ ). This experiment was conducted across of two separate

cohorts of rats and we observed a comparable effect on omissions in both. Across these groups, there were five rats in total that made no choices within one of the blocks of ten free-choice trials following treatment with one or both of the doses, precluding us from analyzing choice data across trial block. To accommodate for this, we averaged the percentage of choices of the large reward across the entire session. This analysis revealed that treatment with the high, but not low dose of quinpirole induced a subtle, but significant reduction in large reward choices ( $F(2,18) = 4.47, p = 0.03$ ; and Dunnett's,  $p < 0.05$ ; Fig. 5E). Quinpirole also increased choice latencies ( $F(2,18) = 4.42, p = 0.01$ ) but did not affect locomotion ( $F(2,18) = 0.14, p = 0.87$ ; Table 2). Collectively, these data suggest that the reduction in risky choice induced by blockade of mOFC D<sub>1</sub> receptors was not driven by a reduction in preference for larger vs smaller rewards. In comparison, excessive stimulation of mOFC D<sub>2</sub> receptors caused a slight reduction in preference for larger rewards, which may have partially contributed to the reduction in risky choice induced by this dose.

**DISCUSSION**

The present findings provide novel insight into how D<sub>1</sub> and D<sub>2</sub> receptors activity within the mOFC modulate action selection in different situations involving reward uncertainty—probabilistic



**Fig. 5** Effects of DA agonist in the mOFC on probabilistic discounting and reward magnitude discrimination. **A** Stimulation of mOFC D<sub>1</sub> receptors did not affect choice for the larger reward (*n* = 13). **B** Excess mOFC D<sub>2</sub> stimulation reduced percentage choice of the large/risky option in the 50% trial block (*n* = 14). **C** Win-stay/lose-shift ratios. Both doses of quinpirole increased lose-shift behavior, and the high dose also reduced win-stay behavior. **D** D<sub>1</sub> blockade with SCH 23390 did not affect preference for larger vs smaller rewards during a reward magnitude discrimination task, left: percent choice of the large reward partitioned over blocks of ten trials; right: percent choice averaged across blocks (demonstrating the lack of a significant main effect of treatment). **E** D<sub>2</sub> stimulation caused a subtle reduction in choice for the large reward option in the reward magnitude discrimination task following treatment with the high, but not low dose of quinpirole. Data are presented as percent choice averaged across the four blocks of free-choice trials to accommodate for increases in trial omissions.

reversal learning and risk/reward decision-making. Pharmacological reduction of mOFC D<sub>1</sub> activity revealed that these receptors aid in optimizing reward seeking, in part by reducing sensitivity to recent nonrewarded actions. Conversely, reducing D<sub>2</sub> activity induced seemingly opposite changes compared to the effects of D<sub>1</sub> antagonism on both probabilistic reversal and discounting, whereas excessive stimulation of D<sub>2</sub>, but not D<sub>1</sub> receptors had deleterious effects across both tasks (see Table 3 for a summary of these results).

#### D<sub>1</sub> receptor modulation of mOFC function

Antagonism of mOFC D<sub>1</sub> receptors perturbed probabilistic reinforcement learning, entailing choice between different actions associated with higher vs lower probabilities of reward. These treatments reduced reversals but also more importantly increased erroneous responding from the initial discrimination of the session indicating a more fundamental impairment in probabilistic choice. mOFC D<sub>1</sub> blockade also reduced preference for larger/risky rewards when animals chose between rewards that differed in terms of magnitude and probability. Moreover, D<sub>1</sub> blockade rendered rats more reliant on recently observable outcomes, as perturbations on both tasks were driven by increased sensitivity to recent negative feedback, in that rats were more likely to shift

**Table 3.** Summary of the effects of D<sub>1</sub> and D<sub>2</sub> blockade and stimulation on probabilistic reversal learning, probabilistic discounting, and reward magnitude discrimination.

	Antagonist		Agonist	
	D <sub>1</sub>	D <sub>2</sub>	D <sub>1</sub>	D <sub>2</sub>
<b>Probabilistic reversal</b>				
Reversals	↓	↑	∅	↓
Errors	↑	↓	∅	∅
Win-stay/Lose-shift	↑ <sub>LS</sub>	↓ <sub>LS</sub>	∅	↑ <sub>LS</sub>
<b>Probabilistic discounting</b>				
Risky choice	↓	↑	∅	↓
Win-stay/Lose-shift	↑ <sub>WS/LS</sub>	∅	∅	↑ <sub>WS/LS</sub>

↓ decrease, ↑ increase, ∅ no change, WS win-stay, LS lose-shift.

response selection after a nonrewarded choice. In this regard, lesions of the primate mOFC impairs probabilistic reinforcement learning that is associated with increased trial-by-trial switching, suggesting this region supports the maintenance of high-value

choices when reward contingencies are probabilistic [11, 35, 36]. Taken together, these findings highlight a key role for mOFC  $D_1$  receptor tone in refining action selection when reward contingencies are probabilistic, promoting more profitable reward seeking by mitigating the impact that probabilistic reward omissions have over subsequent choice.

Regulation of loss sensitivity appears to be a common function of  $D_1$  receptors across multiple cortico-striatal DA terminal regions, as similar effects have been observed following reduction of  $D_1$  receptor activity in the prelimbic region of the mPFC [32] and the nucleus accumbens [37]. Thus,  $D_1$  receptors situated in various nodes within DA decision-making circuitry aid in bridging the gap between rewarded and nonrewarded actions, helping to maintain more profitable choice patterns in the face of negative feedback. This may be related to the proposed function of  $D_1$  receptors in promoting persistent patterns of activity within networks of prefrontal neurons and stabilizing action/outcome representations even in the face of distractors [38–40].

Impairments in probabilistic reversal learning induced by mOFC  $D_1$  antagonism contrasts with the lack of effect of systemic administration of SCH 23390 on a variation of the probabilistic reversal task used here, where trials were self-paced, explicit cues were paired with rewarded choices, and time-out punishments were given for nonrewarded ones [22]. In a similar vein, the reduction in risky choice following mOFC  $D_1$  antagonism was opposite to the increased risky choice and enhanced reward sensitivity induced by inactivation of this region [9]. Note that there have been other reports that administration of dopaminergic drugs can induce different effects on probabilistic choice when administered systemically vs locally in certain terminal regions [32, 37, 41]. Likewise, mOFC inactivation or  $D_1$  antagonism caused opposing effects on effort-related responding, with the former increasing and latter decreasing pursuit of rewards delivered on a progressive ratio schedule [27, 42]. These contrasting effects highlight the importance of distinguishing how DA receptor activity within different nodes of cortico-limbic-striatal circuitry may differentially influence distinct components of reward-seeking behavior.

Stimulation of mOFC  $D_1$  receptors with SKF 81297 had no discernable effects on behavior. Although these null effects should be taken with caution, it is interesting to compare them with a substantial literature showing that increasing  $D_1$  receptor activity within the adjacent medial prefrontal cortex can have either beneficial or detrimental effects on various forms of cognition and executive function [43]. These include attention [33, 44], working memory [33, 45], and risk/reward decision-making [32]. With respect to the present study, even though mesocortical  $D_1$  receptors are thought to be relatively unoccupied under basal conditions, the fact that mOFC  $D_1$  receptor blockade did alter probabilistic learning and discounting suggest that these receptors were activated during performance of these tasks. The lack of effect of additional  $D_1$  receptor stimulation under these conditions suggests that their activation by endogenous DA was maximal, and that these functions were relatively impervious to additional  $D_1$  receptor stimulation. In this regard, it is notable that other functions, such as set-shifting, are also unaffected by increased  $D_1$  receptor tone in the medial prefrontal cortex, even though blockade of these receptors impairs this form of cognitive flexibility [43]. This combination of findings suggests that the principles of operation underlying DA modulation of frontal lobe function can vary across different cognitive processes and prefrontal regions.

#### $D_2$ receptor modulation of mOFC function

Blockade of mOFC  $D_2$  receptors induced effects that were seemingly opposite to those induced by  $D_1$  receptor blockade. Intra-mOFC infusions of eticlopride increased reversals completed, but this was once again accompanied by a reduction in errors and

lose-shift behavior during the early phases of the task.  $D_2$  blockade facilitated identification of the more profitable option and promoted a strategy for repeating successful decisions despite losses that occurred. Thus, whereas  $D_1$  blockade induced a “choice-switching” phenotype where rats had difficulty stabilizing choice bias toward the more profitable action-outcome association,  $D_2$  blockade induced a “choice-maintenance” phenotype, where rats exploited the more valuable option during the first discrimination of the task. It is important to note that while we did see changes in reversal performance, the fact that both  $D_1$  and  $D_2$  antagonism altered errors and sensitivity to recent losses during the initial discrimination suggests these effects may be more attributable to changes on probabilistic learning, rather than flexibility per se. In a similar vein, blocking mOFC  $D_2$  receptors increased risky choice, again orienting rats toward recently observable feedback, as rats showed a trend to follow a risky win with another risky choice. Taken together,  $D_2$  blockade in the mOFC led to a stronger maintenance of profitable choice in both instances, suggesting that normal mOFC  $D_2$  receptor tone serves to mitigate strong choice biases, specifically in response to probabilistic wins and losses. This would be particularly advantageous when unexpected changes in action-outcome contingencies occur, to support behavioral alterations in response to changes in these contingencies. This may be related to the proposed function of  $D_2$  receptors in biasing prefrontal network activity away from robustness, allowing many items to be represented and compared simultaneously. In turn, this would facilitate response flexibility, where it is advantageous to sample and compare different options ([38] and reviewed in [39, 40]).

The idea that  $D_1$  and  $D_2$  receptors within a particular brain region mediate distinct patterns of behavior is certainly not novel. Previously, we have shown that  $D_1$  or  $D_2$  receptors in the adjacent prelimbic prefrontal cortex can promote different aspects of choice behavior via modulation of distinct networks of neurons that interface with different subcortical targets [20]. In this regard, prefrontal  $D_1$  receptors serve to reinforce actions yielding larger rewards via a network that interfaces with the nucleus accumbens, while  $D_2$  receptors facilitate flexibility in decision biases via actions on networks that interface with the basolateral amygdala [20]. One particularly interesting feature of the present findings is that mOFC  $D_1$  receptor antagonism induced an effect similar to mOFC inactivation on the probabilistic reversal task, whereas on the probabilistic discounting task, it was  $D_2$  receptor blockade that produced an effect comparable to mOFC inactivation. Thus, it is plausible that, like DA modulation within the medial prefrontal cortex, dissociable networks of mOFC neurons may be differentially modulated by  $D_1$  or  $D_2$  receptors that are recruited in different manners under the two task conditions assayed here. Although both tasks entail goal-directed behavior involving reward uncertainty, there are fundamental differences in the types of information that are processed to guide action selection. The probabilistic reversal task requires a choice between different actions that may lead to rewards of a fixed magnitude. Conversely, the probabilistic discounting task requires a choice between rewards of differing magnitudes, the larger of which incurs an uncertainty cost that can reduce its utility. In light of these considerations, it seems reasonable to propose that different dopaminergic mechanisms may underlie mOFC mediation of choosing between different actions that may yield reward versus choosing between rewards that differ in value.

In contrast to the effects of  $D_2$  blockade, excessive  $D_2$  receptor stimulation with quinpirole disrupted choice behavior. During probabilistic reversals, quinpirole reduced the number of reversals completed. This is in keeping with previous reports that systemic administration of quinpirole impairs reversal performance and altered negative feedback learning [22, 23]. Quinpirole also reduced risky choice during probabilistic discounting, specifically during the 50% trial block, accompanied by increased lose-shift

behavior. It is important to note that the high dose of quinpirole also reduced preference for larger vs smaller rewards when both options were certain, but this effect was both subtle in magnitude, and not apparent at the low dose which was still effective at altering probabilistic choice. This reduction of preference for larger rewards may have contributed to the reduced win–stay behavior induced by the high dose of quinpirole during probabilistic discounting. This combination of findings suggests that excessive D<sub>2</sub> receptor activity within the mOFC may reduce choice biases toward larger rewards, an effect that is more prominent when these rewards are uncertain. It is notable that this effect contrasts to a certain degree with what has been observed following D<sub>2</sub> stimulation in the prelimbic cortex, which blunted reward sensitivity on risk/reward decision-making and led to more static patterns of choice [32]. This adds additional credence to the notion that mesocortical DA transmission may subserve complementary yet distinct functions within different frontal lobes regions

In addition to having detrimental effects on choice behavior, excessive D<sub>2</sub> stimulation also reduced task engagement, as indexed by increased trial omissions in all three assays used here. Notably, similar effects have been reported following pharmacological increases in D<sub>2</sub> receptor activity in the lateral OFC, as quinpirole infusions in this region increased omissions during the five-choice serial reaction time task [46]. This implicates D<sub>2</sub> receptors across the whole OFC in regulating a more basic motivation or attention for reward, and raises the possibility that they may modulate a population of mOFC neurons that project to the striatum to invigorate reward seeking.

## CONCLUSIONS

The findings reported here reveal previously uncharacterized roles for DA within the mOFC in biasing goal-directed and reward-related behavior in the face of probabilistic rewards. mOFC D<sub>1</sub> and D<sub>2</sub> receptors play dissociable and opposing roles in biasing cost/benefit analyses when choosing between rewards that differ in terms of magnitude and uncertainty and using reward history to guide choice toward options that are more likely to yield rewards. It is important to note that much of what we know of how DA may modulate prefrontal network states comes from studies of the prelimbic region of the mPFC. In comparison, there have been substantially fewer studies probing how DA may regulate neural activity within the mOFC. Nevertheless, given the similar cellular and neurochemical make up of these two frontal lobe regions, it is plausible that while prefrontal vs mOFC DA transmission appear to play distinct roles in guiding reward-related action selection, the basic principles governing mesocortical DA modulation of network states may be consistent across these two regions.

These findings represent a first step in understanding how the mOFC, and more specifically, DA transmission in this region contributes as part of the broader neural circuitry involved in biasing adaptive goal-directed action in males, although exploring whether these functions generalize to females remains an important topic for future research. As such, these findings may have important implications for understanding how D<sub>1</sub> and D<sub>2</sub> receptors may regulate both normal and abnormal functions supported by the OFC. Mesocortical DA contributes to cognitive dysfunction in diseases such as schizophrenia, Parkinson's, and substance abuse disorders, all of which are associated with changes in OFC structure [47–52]. Future studies aiming to clarify how mOFC DA regulates both normal and abnormal cognitive functions will need to consider the balance between the opposing actions of D<sub>1</sub> and D<sub>2</sub>, as well as the downstream targets to which these DA receptors signal.

## FUNDING AND DISCLOSURE

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-04295; probabilistic reversal experiments) and a Project Grant from the Canadian Institutes of Health Research (PJT-162444—risk/reward decision-making experiments) to SBF and an NSERC Graduate Fellowship to NLJ. The authors declare no competing interests.

## AUTHOR CONTRIBUTIONS

NLJ and YTL acquired the data. NLJ and SBF designed the study, analyzed the data, and wrote the manuscript.

## ADDITIONAL INFORMATION

**Supplementary Information** accompanies this paper at (<https://doi.org/10.1038/s41386-020-00931-1>).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

1. Izquierdo A. Functional heterogeneity within rat orbitofrontal cortex in reward learning and decision making. *J Neurosci.* 2017;37:10529.
2. Carmichael ST, Price JL. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol.* 1996;371:179–207.
3. Ongür D, Price JL. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex.* 2000;10:206–19.
4. Heilbronner SR, Rodriguez-Romaguera J, Quirk GJ, Groenewegen HJ, Haber SN. Circuit-based corticostriatal homologies between rat and primate. *Biol Psychiatry.* 2016;80:509–21.
5. Hoover WB, Vertes RP. Projections of the medial orbital and ventral orbital cortex in the rat. *J Comp Neurol.* 2011;519:3766–801.
6. Malvaez M, Shieh C, Murphy MD, Greenfield VY, Wassum KM. Distinct cortical–amygdala projections drive reward value encoding and retrieval. *Nat Neurosci.* 2019;22:762–9.
7. Dalton GL, Wang NY, Phillips AG, Floresco SB. Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. *J Neurosci.* 2016;36:1996–2006.
8. Bradfield LA, Dezfouli A, Van Holstein M, Chieng B, Balleine BW. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. *Neuron.* 2015;88:1268–80.
9. Stopper CM, Green EB, Floresco SB. Selective involvement by the medial orbitofrontal cortex in biasing risky, but not impulsive, choice. *Cereb Cortex.* 2014;24:154–62.
10. Balleine BW, Leung BK, Ostlund SB. The orbitofrontal cortex, predicted value, and choice. *Ann N Y Acad Sci.* 2011;1239:43–50.
11. Noonan MP, Kolling N, Walton ME, Rushworth MFS. Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *Eur J Neurosci.* 2012;35:997–1010.
12. Gourley SL, Zimmermann KS, Allen AG, Taylor JR. The medial orbitofrontal cortex regulates sensit outcome value. *J Neurosci.* 2016;36:4600–13.
13. Sharpe MJ, Wikenheiser AM, Niv Y, Schoenbaum G. The state of the orbitofrontal cortex. *Neuron.* 2015;88:1075–7.
14. Dalton GL, Phillips AG, Floresco SB. Preferential involvement by nucleus accumbens shell in mediating probabilistic learning and reversal shifts. *J Neurosci.* 2014;34:4618–26.
15. Hall-McMaster S, Millar J, Ruan M, Ward RD. Medial orbitofrontal cortex modulates associative learning between environmental cues and reward probability. *Behav Neurosci.* 2016;131:1–10.
16. Clark L, Bechara A, Damasio H, Aitken MRF, Sahakian BJ, Robbins TW. Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain.* 2008;131:1311–22.
17. Hervig ME, Fiddian L, Piilgaard L, Bo T, Knudsen C, Olesen SF, et al. Dissociable and paradoxical roles of rat medial and lateral orbitofrontal cortex in visual serial reversal learning. *Cereb Cortex.* 2020;30:1016–29.
18. Floresco SB, Magyar O, Ghods-Sharifi S, Vexelman C, Tse MTL. Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology.* 2006;31:297–309.

19. Seamans JK, Floresco SB, Phillips AG. D1 receptor modulation of hippocampal-prefrontal cortical circuits integrating spatial memory with executive functions in the rat. *J Neurosci*. 1998;18:1613–21.
20. Jenni NL, Larkin JD, Floresco SB. Prefrontal dopamine D1 and D2 receptors regulate dissociable aspects of risk/reward decision-making via distinct ventral striatal and amygdalar circuits. *J Neurosci*. 2017;37:6200–13.
21. St. Onge JR, Stopper CM, Zahm DS, Floresco SB. Separate prefrontal-subcortical circuits mediate different components of risk-based decision making. *J Neurosci*. 2012;32:2886–99.
22. Verharen JPH, Adan RAH, Vanderschuren LJM. Differential contributions of striatal dopamine D1 and D2 receptors to component processes of value-based decision making. *Neuropsychopharmacology*. 2019;44:2195–204.
23. Alsio J, Phillips BU, Sala-Bayo J, Nilsson SRO, Calafat-Pla TC, Rizwand A, et al. Dopamine D2-like receptor stimulation blocks negative feedback in visual and spatial reversal learning in the rat: behavioural and computational evidence. *Psychopharmacology*. 2019;236:2307–23.
24. Oades RD, Halliday GM. Ventral tegmental (A10) system: neurobiology. 1. Anatomy and connectivity. *Brain Res Rev*. 1987;12:117–65.
25. Tomm RJ, Tse MT, Tobiansky DJ, Schweitzer HR, Soma KK, Floresco SB. Effects of aging on executive functioning and mesocorticolimbic dopamine markers in male Fischer 344 × brown Norway rats. *Neurobiol Aging*. 2018;72:134–46.
26. Cosme CV, Gutman AL, Worth WR, Lalumiere RT. D1, but not D2, receptor blockade within the infralimbic and medial orbitofrontal cortex impairs cocaine seeking in a region-specific manner. *Addict Biol*. 2016;23:16–27.
27. Münster A, Sommer S, Hauber W. Dopamine D1 receptors in the medial orbitofrontal cortex support effort-related responding in rats. *Eur Neuropsychopharmacol*. 2020;32:136–41.
28. Orsini CA, Setlow B. Sex differences in animal models decision making. *J Neurosci Res*. 2017;269:260–9.
29. Islas-preciado D, Wainwright SR, Sniegocki J, Liebllich SE, Yagi S, Floresco SB, et al. Hormones and behavior risk-based decision making in rats: modulation by sex and amphetamine. *Horm Behav*. 2020;125:104815.
30. Bari A, Theobald DE, Caprioli D, Mar AC, Aidoo-Micah A, Dalley JW, et al. Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology*. 2010;35:1290–301.
31. St. Onge JR, Floresco SB. Prefrontal cortical contribution to risk-based decision making. *Cereb Cortex*. 2010;20:1816–28.
32. St. Onge JR, Abhari H, Floresco SB. Dissociable contributions by prefrontal D1 and D2 receptors to risk-based decision making. *J Neurosci*. 2011;31:8625–33.
33. Chudasama Y, Robbins TW. Dopaminergic modulation of visual attention and working memory in the rodent prefrontal cortex. *Neuropsychopharmacology*. 2004;29:1628–36.
34. Verharen JPH, den Ouden HEM, Adan RAH, Vanderschuren LJM. Modulation of value-based decision making behavior by subregions of the rat prefrontal cortex. *Psychopharmacology*. 2020;237:1267–80.
35. Clarke HF, Cardinal RN, Rygula R, Hong YT, Fryer TD, Sawiak SJ, et al. Dopamine alters behavior via changes in reinforcement sensitivity. *J Neurosci*. 2014;34:7663–76.
36. Noonan MP, Walton ME, Behrens TEJ, Sallet J, Buckley MJ, Rushworth MFS. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci USA*. 2010;107:20547–52.
37. Stopper CM, Khayambashi S, Floresco SB. Receptor-specific modulation of risk-based decision making by nucleus accumbens dopamine. *Neuropsychopharmacology*. 2013;38:715–28.
38. Seamans JK, Yang CR. The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol*. 2004;74:1–57.
39. Durstewitz D, Seamans JK. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-O-methyltransferase genotypes and schizophrenia. *Biol Psychiatry*. 2008;64:739–49.
40. Cools R, D'Esposito M. Inverted-U-shaped dopamine actions on human working memory and cognitive control. *Biol Psychiatry*. 2011;69:e113–125.
41. St. Onge JR, Floresco SB. Dopaminergic modulation of risk-based decision making. *Neuropsychopharmacology*. 2009;34:681–97.
42. Münster A, Hauber W. Medial orbitofrontal cortex mediates effort-related responding in rats. *Cereb Cortex*. 2018;28:4379–89.
43. Floresco SB. Prefrontal dopamine and behavioral flexibility: Shifting from an 'inverted-U' toward a family of functions. *Front Neurosci*. 2013;7:1–12.
44. Granon S, Passetti F, Thomas KL, Dalley JW, Everitt BJ, Robbins TW. Enhanced and impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. *J Neurosci*. 2000;20:1208–15.
45. Floresco SB, Phillips AG. Delay-dependent modulation of memory retrieval by infusion of a dopamine D1 agonist into the rat medial prefrontal cortex. *Behav Neurosci*. 2001;115:934–9.
46. Winstanley CA, Zeeb FD, Bedard A, Fu K, Lai B, Steele C, et al. Dopaminergic modulation of the orbitofrontal cortex affects attention, motivation and impulsive responding in rats performing the five-choice serial reaction time task. *Behav Brain Res*. 2010;210:263–72.
47. Mladinov M, Sedmak G, Fuller HR, Babić Leko M, Mayer D, Kirincich J, et al. Gene expression profiling of the dorsolateral and orbital prefrontal cortex in schizophrenia. *Transl Neurosci*. 2016;7:139–50.
48. Moorman DE. The role of the orbitofrontal cortex in alcohol use, abuse, and dependence. *Prog Neuro-Psychopharmacology*. *Biol Psychiatry*. 2018;87:85–107.
49. Van Eimeren T, Ballanger B, Pellecchia G, Janis MM, Lang AE, Strafella AP. Dopamine agonists diminish value sensitivity of the orbitofrontal cortex: a trigger for pathological gambling in Parkinson's disease? *Neuropsychopharmacology*. 2009;34:2758–66.
50. Robbins TW, Cools R. Cognitive deficits in Parkinson's disease: a cognitive neuroscience perspective. *Mov Disord*. 2014;29:597–607.
51. Tanabe J, Tregellas JR, Dalwani M, Thompson L, Owens E, Crowley T, et al. Medial orbitofrontal cortex gray matter is reduced in abstinent substance-dependent individuals. *Biol Psychiatry*. 2009;65:160–4.
52. Volkow ND, Chang L, Wang G-J, Fowler JS, Ding Y-S, Sedler M, et al. Low level of brain dopamine D2 receptors in methamphetamine abusers: association with metabolism in the orbitofrontal cortex in methamphetamine abusers: association with metabolism in the orbitofrontal cortex. *Am J Psychiatry*. 2001;158:2015–21.