Cell Reports
Article



# A Role for Dopamine-Mediated Learning in the Pathophysiology and Treatment of Parkinson's Disease

Jeff A. Beeler,[1,5,*] Michael J. Frank,[4,5] John McDaid,[2,5] Erin Alexander,[1] Susie Turkson,[1] Maria Sol Bernandez,[1] Daniel S. McGehee,[2,3] and Xiaoxi Zhuang[1,3]
[1]Department of Neurobiology
[2]Department of Anesthesia and Critical Care
[3]Committee on Neurobiology
The University of Chicago, Chicago, IL 60637, USA
[4]Department of Cognitive, Linguistic, and Psychological Sciences, and Department of Psychiatry and Human Behavior and Brown Institute for Brain Science, Brown University, Providence, RI 02912, USA
[5]These authors contributed equally to this work
*Correspondence: jabeeler@uchicago.edu
http://dx.doi.org/10.1016/j.celrep.2012.11.014

## SUMMARY

**Dopamine contributes to corticostriatal plasticity and motor learning. Dopamine denervation profoundly alters motor performance, as in Parkinson's disease (PD); however, the extent to which these symptoms reflect impaired motor learning is unknown. Here, we demonstrate a D2 receptor blockade-induced aberrant learning that impedes future motor performance when dopamine signaling is restored, an effect diminished by coadministration of adenosine antagonists during blockade. We hypothesize that an inappropriate corticostriatal potentiation in striatopallidal cells of the indirect pathway underlies aberrant learning. We demonstrate synaptic potentiation in striatopallidal neurons induced by D2 blockade and diminished by application of an adenosine antagonist, consistent with behavioral observations. A neurocomputational model of the basal ganglia recapitulates the behavioral pattern and further links aberrant learning to plasticity in the indirect pathway. Thus, D2-mediated aberrant learning may contribute to motor deficits in PD, suggesting new avenues for the development of therapeutics.**

## INTRODUCTION

Corticostriatal plasticity has been directly linked to motor learning and performance (Costa et al., 2004; Yin et al., 2009; Jin and Costa, 2010). The dorsolateral striatum (posterior putamen in primates)—the region most prominently affected in Parkinson's disease (PD) (Bernheimer et al., 1973; Hornykiewicz, 2001)— has been associated with the automization of behavior (Miyachi et al., 2002; Costa et al., 2004; Poldrack et al., 2005; Puttemans et al., 2005; Doyon et al., 2009; Yin et al., 2009; Jin and Costa, 2010) and habit (Bernheimer et al., 1973; Hornykiewicz, 2001; Tang et al., 2007; Graybiel, 2008; Yin et al., 2009; Balleine and O'Doherty, 2010), providing a substrate for generating rapid and efficient behavioral responses without cognitive deliberation and planning.

Dopamine denervation induces abnormal corticostriatal plasticity (Calabresi et al., 1997; Picconi et al., 2003; Kreitzer and Malenka, 2007; Shen et al., 2008; Peterson et al., 2012), though the role this plays in the symptoms, progression, and treatment of the PD has not been established. We recently proposed that altered plasticity, specifically inappropriate LTP in striatopallidal medium-spiny neurons (MSNs), gives rise to an aberrant learning process that contributes to the symptoms and progression of PD by inverting basal ganglia optimization of behavior (Wiecki and Frank, 2010; Beeler, 2011). Computational models suggest an interaction between dopamine's effects on MSN activity and corticostriatal synaptic plasticity—performance and learning, respectively— within striatal D1- and D2-expressing cells of the direct and indirect pathways (Frank et al., 2004; Bódi et al., 2009; Palminteri et al., 2009; Wiecki and Frank, 2010). To the degree that the mechanisms of abnormal corticostriatal plasticity are dissociable from those mediating dopamine's direct performance effects, they represent a target for novel therapeutics. Remediating abnormal plasticity and aberrant learning may be a significant but unrecognized component of current drug therapies and underlie the poorly understood but important long-duration response (LDR) observed in L-DOPA treatment (Beeler et al., 2010; Beeler, 2011).

Both the aberrant learning hypothesis and neurocomputational models point to critical interactive effects between dopamine-mediated performance and learning and make specific predictions:

(1) Dopamine depletion or receptor blockade, in addition to direct performance effects, will result in inhibitory learning in the indirect pathway that will impair future performance and learning even when dopamine signaling is restored.
(2) In animals that acquired the task under healthy dopamine conditions, dopamine blockade should induce a

progressive decline in performance, reflecting an aberrant learning process that will impair future recovery.

(3) If the above effects are due to induction of aberrant potentiation in the D2 pathway, dopamine blockade should induce potentiation in striatopallidal MSNs that is reversed by agents known to disrupt LTP in this pathway.

(4) Disrupting LTP in the indirect pathway should be protective when administered during dopamine blockade by preventing aberrant learning but impede recovery when administered after aberrant learning by impairing relearning.

Here, we test these predictions in a mouse model of motor learning and concurrently test whether the empirically observed effects will emerge in an a priori computational model of basal ganglia function.

## RESULTS

### Blockade of D1 and D2 Signaling Induce Aberrant Learning Independent of Direct Effects on Performance

To reversibly mimic dopamine denervation, we administered a cocktail of D1 and D2 dopamine receptor antagonists to block dopamine signaling in the direct and indirect pathways. The cocktail was administered either during the initial acquisition or after the establishment of a striatal-dependent motor skill. When a dopamine antagonist cocktail is administered during initial training, the cocktail dramatically impaired rotarod performance (treatment main effect, $F_{(1,12)} = 34.5$, $p < 0.001$) with no apparent evidence of learning across the 5 training days (Figure 1A, left). However, when the mice returned to the rotarod 72 hr after the last cocktail administration, their performance remains degraded and improves only gradually over 10 days (Figure 1A, treatment x session interaction, $F_{(1,12)} = 23.6$, $p < 0.001$), which contrasts markedly with the 1–2 days required for naive mice to reach asymptotic performance (Figure 1A, left; comparing initial 3 days naive versus relearning, treatment x session interaction, $F_{(1,12)} = 11.29$, $p < 0.01$). The dramatically retarded reacquisition following initial training under cocktail suggests that an aberrant learning process occurred during the cocktail training. This was not a residual effect of the cocktail treatment itself: mice administered the cocktail on the same schedule as the trained mice but without rotarod training showed no impairment in subsequent acquisition (Figure 1A, right), indicating the diminished performance of cocktail trained mice is experience dependent.

We further tested if cocktail together with rotarod training impaired the animals on other tasks. In a new group, we administered cocktail and provided rotarod training but following the 72 hr break, tested their open field (OF) activity and treadmill performance instead of rotarod. We compared these mice to a naive control group and a group administered eticlopride acutely as a reference for impaired D2 function. Although decreased D2 activation (i.e., eticlopride) greatly reduces OF activity, the cocktail/rotarod group does not show decreased activity (Figure 1C; $F_{(2,18)} = 39.12$, $p < .001$). On the first two days, they show increased activity compared with the naive mice. The reason for this is unclear but inconsistent with decreased D2 function. Notably, the cocktail/rotarod group had a week of handling and exercise that may have increased their comfort level and decreased behavioral inhibition. Consistent with this possibility, they show greater rearing/vertical time, consistent with greater exploratory behavior (Figure 1D; $F_{(2,18)} = 16.46$, $p < 0.001$). In a second motor performance test, performance on the treadmill was indistinguishable between naive and cocktail/rotarod groups, in contrast to greatly reduced performance as a consequence of reduced D2 signaling (Figure 1E; group, $F_{(1,22)} = 28.3$, $p < 0.001$; naive versus cocktail trained, $F_{(1,14)} = 2.25$, $p = .15$). Together, these data suggest that the training under D1/D2 blockade resulted in an experience-dependent aberrant motor learning that task-specifically impaired subsequent performance and learning.

When the antagonist cocktail was administered after the motor skill was established (in a separate group of mice), we observed an aberrant learning effect as well. The cocktail results in an immediate impairment of performance (treatment main effect, $F_{(1,10)} = 58.7$, $p < 0.001$) but also a further decline across the initial couple of sessions (Figure 1B, middle; treatment group only, session effect, $F_{(1,5)} = 7.2$, $p = 0.013$). Upon cessation of cocktail treatment, the mice did not return immediately to their prior asymptotic performance. Instead, they showed a gradual return to previous levels of performance (Figure 1B, right; during recovery phase: prior treatment $F_{(1,10)} = 7.2$, $p = 0.023$, prior treatment x session, $F_{(1,10)} = 5.8$, $p = 0.036$). These data suggest that the cocktail not only impairs performance but when combined with experience on the motor task, also alters the underlying learning that supported the previously established skill, necessitating a relearning phase.

The cocktail blocks both the D1 and D2 receptors. Previous reports suggest that aberrant learning occurs primarily in the D2 pathway (Wiecki et al., 2009; Beeler et al., 2010). To independently test the contribution of each receptor subtype, we tested the behavioral consequences of administering eticlopride (D2 selective antagonist) or SCH23390 (D1 selective antagonist), either during acquisition or after asymptotic performance, as above.

### Blockade of D2 but Not D1 Receptors during Acquisition Induces Aberrant Learning that Impairs Subsequent Recovery

Both the D1 and D2 antagonists impair initial acquisition (Figures 2A and 2B, left; acquisition only, treatment effect, $F_{(7,23)} = 7.78$, $p < 0.001$) in a dose-dependent manner (SCH23390 dose, $F_{(1,18)} = 19.87$, $p < 0.001$; eticlopride dose, $F_{(1,13)} = 5.39$, $p < 0.05$), though not to the same degree as the cocktail (Figure 1A, left). At higher doses, D1 and D2 antagonists impaired performance to a similar degree. However, a significantly different pattern of recovery emerged during the drug-free phase depending on whether D1 or D2 receptors were blocked during acquisition (drug x session during recovery, $F_{(1,25)} = 16.09$, $p < 0.001$). Blockade of D1 has a weak effect on subsequent recovery (Figures 2A and 2C; SCH23390 group during recovery, dose, $F_{(1,18)} = 4.14$, $p = 0.056$). Indeed, we observed an immediate, discontinuous jump to better performance posttreatment in D1-treated mice with little subsequent improvement (Figures 2A and 2C; SCH23390 recovery, sessions $F_{(1,18)} = 0.79$,

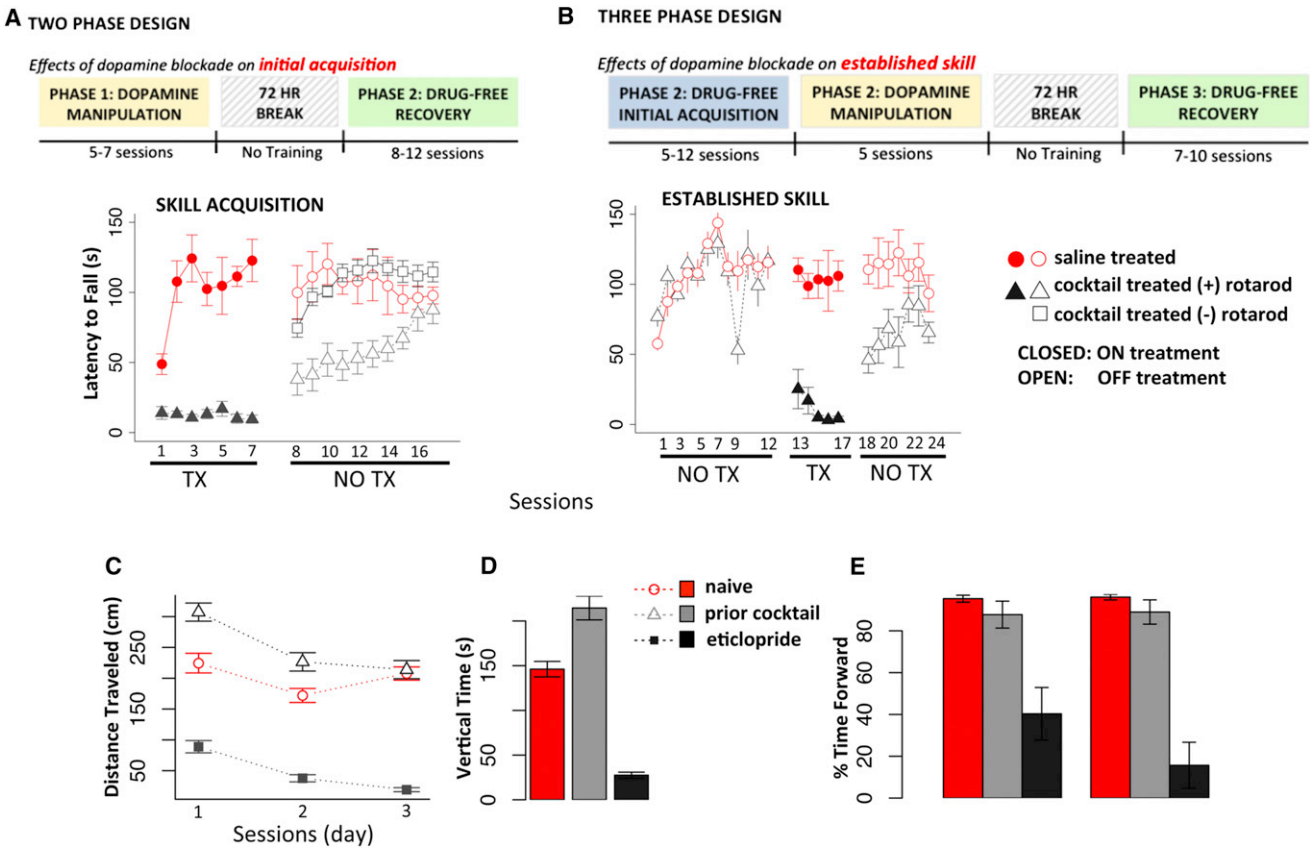


**Figure 1. Dissociating Learning and Performance Effects of Dopamine Receptor Blockade on Acquisition and Maintenance of a Motor Skill**

(A and B) To dissociate learning and performance effects, we used a multiphase rotarod design in which initial learning under either a dopamine-normal or dopamine-impaired condition is paired with a subsequent testing phase in the opposite condition. A two- and three-phase design (shown above graphs) was used to assess the effects of dopamine-receptor blockade on initial acquisition (two-phase design) or continued performance of an established skill (three-phase design), respectively. Each daily session consisted of five trials. In all figures, session means averaged across the five trials are reported. Graphs show latency to fall in wild-type C57BL/6 mice administered either a cocktail of dopamine antagonists (0.16 mg/kg eticlopride + 0.1 mg/kg SCH23390, filled gray triangles) or saline (filled red circles) during either (A) initial acquisition of rotarod performance or (B) after performance is established through 12 days prior training. A control group was administered the antagonist cocktail and returned to their homecage without rotarod training and then subsequently tested without drug (A, open gray squares). Each point represents the mean of five trials during daily sessions. Between each treatment phase, there was a 72 hr break without training.

(C) Distance traveled in the open field (OF) for three 45 min sessions on three consecutive days for a group that previously received cocktail administration and rotarod training (open triangles), a naive group (red circles), and a group administered 0.16 mg/kg eticlopride 30 min prior to testing.

(D) Average rearing/vertical time across all three OF sessions.

(E) Average percent of time mice (same groups as in D and E) remained in the forward two-thirds of track on treadmill during two 20 s test trials at 15 and 20 cm/s.

(A) n = 7 (homecage controls, open square, n = 5); (B) n = 6; (C and D) n = 7; and (E) n = 8. Error bars, SEM.

p = 0.38). For highest doses of D1 blockade during acquisition, some residual learning appears to have occurred during recovery (sessions x dose, $F_{(1,18)} = 3.37$, p = 0.082). In contrast, D2 blockade during initial acquisition results in poor initial performance during recovery and only gradual relearning and improvement (Figures 2B and 2D, right; eticlopride group during recovery, sessions $F_{(1,13)} = 21.4$, $p < 0.001$; dose, not significant (N.S.); dose x session, $F_{(1,13)} = 3.36$, p = 0.08). The mild relearning observed at high doses of D1 blockade may reflect a reduction in receptor selectivity at higher doses or indicate that sufficient interference with D1 can also induce mild aberrant learning. Nonetheless, in general, D1 blockade impaired performance with minimal effects on subsequent performance, suggesting a limited role in aberrant learning. In contrast, D2 blockade significantly impairs future performance and appears to require a gradual relearning process. These data suggest that blockade of D2 induces an aberrant learning during initial acquisition that delays or hinders subsequent appropriate learning.

When the antagonists are instead applied after asymptotic performance is established, a difference in the pattern of impairment and recovery emerges. D1 blockade results in an immediate, dose-dependent decrement in performance that is constant across drug administration sessions (Figures 3A and 3C treatment phase; dose, $F_{(1,18)} = 27.1$, $p < 0.001$, session and dose X session, N.S.). Upon cessation of the antagonist, performance in mice administered a D1 antagonist returns to asymptotic performance immediately, showing no dose-dependent effects or relearning (Figures 3A and 3C, recovery phase;

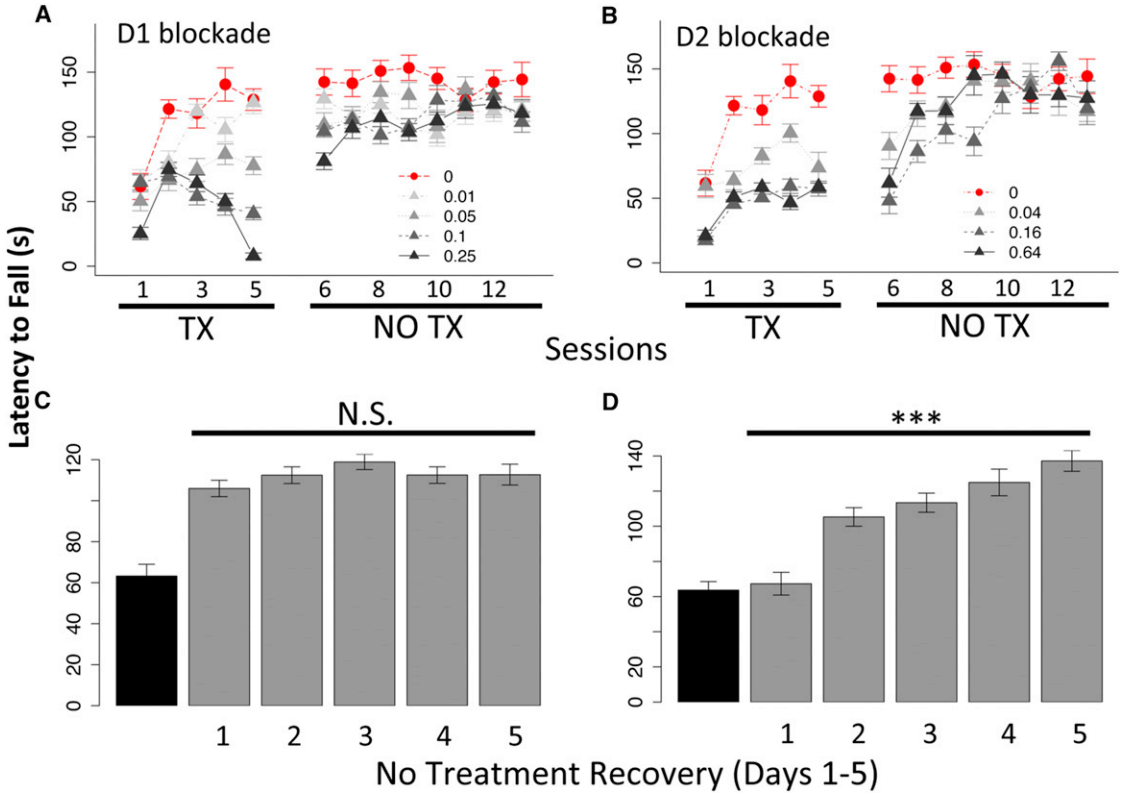


**Figure 2. Effect of D1 or D2 Antagonism on Acquisition and Subsequent Drug-free Recovery**

Latency to fall in wild-type C57BL/6 mice administered an antagonist of either (A) D1 (SCH23390) or (B) D2 (Eticlopride) during 5 days of initial acquisition of rotarod performance and subsequent nondrug performance. Red traces indicate saline controls (shown in both plots). Bar graphs show the average of the last 3 days of drug training (black bars), followed by the first 5 days of drug-free recovery averaged across all doses for (C) D1 and (D) D2 blockade. Each point represents the average of five trials during daily sessions. A 72 hr break occurred between treatment and nontreatment phases. n = 4/dose, ***$p < 0.001$. Error bars, SEM.

dose, $F_{(1,18)} = 0.81$, $p = 0.38$, session and dose X session, N.S.). In contrast, D2 blockade, though it also causes an immediate performance decrement, also induces gradual deterioration across treatment sessions (treatment phase, dose main effect, $F_{(1,14)} = 5.9$, $p < 0.05$, dose X session $F_{(1,14)} = 6.3$, $p < 0.05$). During recovery from D2 blockade, unlike D1 blockade, we observe a continued effect of dose (Figures 3B and 3D, recovery phase, dose, $F_{(1,14)} = 6.29$, $p < 0.05$, session and dose X session, N.S.) and what appears to be a more gradual recovery compared to recovery from D1 blockade, though this latter observation was not statistically significant (drug X session, $F_{(1,15)} = 0.77$, $p = 0.47$).

Though D2 blockade during initial acquisition clearly impairs recovery (Figures 2B and 2D), the degree to which D2 blockade of an established skill impairs subsequent recovery is less clear (Figures 3B and 3D). This could arise because putative aberrant learning may depend on how established the skill is to begin with. Additionally, the effect of D2 blockade (Figure 3D) does not appear to titrate as clearly as D1 blockade (Figure 3C); the lowest dose had no apparent effect (comparing saline and 0.04, dose and dose X session, N.S.) and the 0.16 and 0.64 effect were similar (comparing 0.16 and 0.64, dose and dose X session, N.S.). Consequently, we conducted two additional experiments.

In the first experiment, we provided a longer initial training period (12 days, similar to Figure 1B) and directly compare the effects of D1 and D2 blockade using the doses used in the cocktail (Figure 3E). The different treatments yield significantly different performance effects (treatment, $F_{(1,15)} = 13.9$, $p < 0.001$; treatment x session, $F_{(1,15)} = 3.68$, $p < 0.05$) and significantly different recoveries (recovery, prior treatment effect, $F_{(1,15)} = 9.53$, $p < 0.01$). As observed previously, D1 blockade results in an immediate decrement that is constant across the treatment days (SCH23390 treatment effect, $F_{(1,10)} = 11.45$, $p < 0.01$; session, N.S.; session x treatment, N.S.), with no effects on subsequent drug-free recovery (Figure 3E, SCH23390 prior treatment effect, $F_{(1,10)} = 0.124$, $p = 0.73$). In contrast, D2 blockade induces a gradual deterioration across treatment days (eticlopride treatment effect, $F_{(1,10)} = 20.2$, $p = < 0.01$; treatment x session, $F_{(1,10)} = 4.71$, $p = 0.055$) and impaired performance and slowed recovery during drug-free recovery (Figure 3E, eticlopride prior treatment effect, $F_{(1,10)} = 21.1$, $p < 0.001$; session, $F_{(1,5)} = 5.6$, $p = 0.06$).

In the second experiment, we varied the number of training days prior to administering the D2 blockade (Figure 3F, 12 day group same as 3E). We observe the same pattern as in Figure 3E with no significant difference in either treatment or recovery as

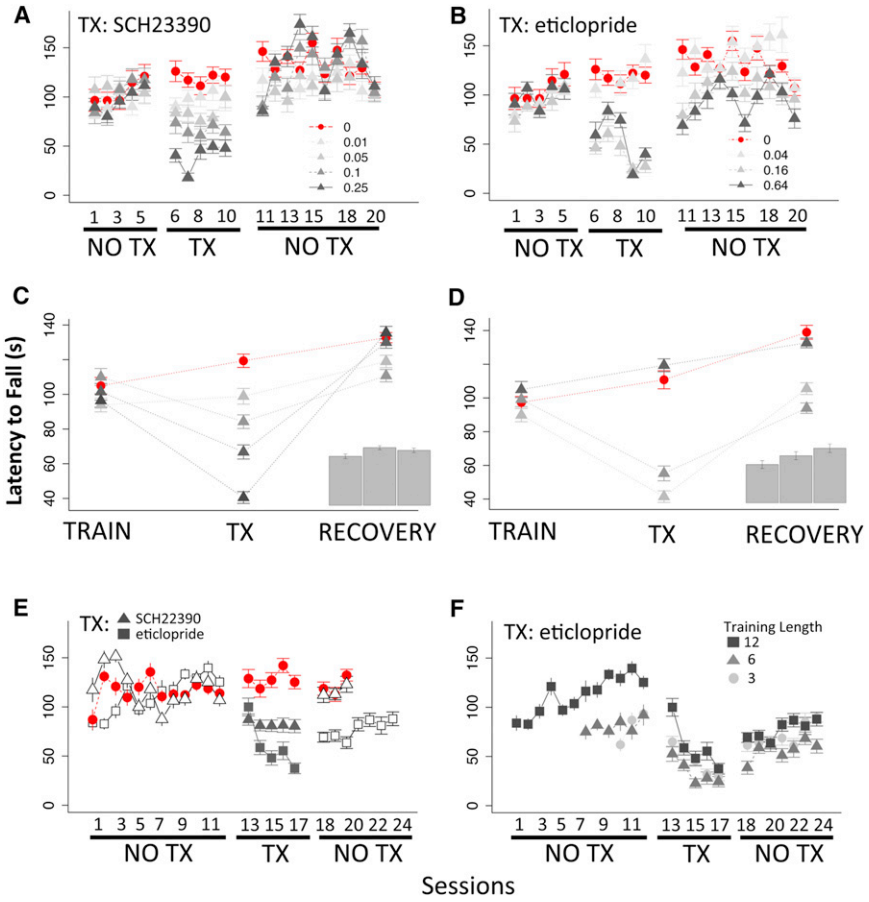


**Figure 3. Effect of D1 or D2 Antagonism on Established Performance**

(A and B) Latency to fall in wild-type C57BL/6 mice administered an antagonist of either (A) D1 (SCH23390) or (B) D2 (Eticlopride) following 5 days of initial training under nondrug conditions and subsequent recovery. Each point represents the average of five trials during daily sessions.

(C and D) Average performance across sessions in each phase plotted by dose for (C) D1 and (D) D2 blockade. Bar graph insets show first three recovery days averaged across all doses of SCH23390 and for 0.16 and 0.64 mg/kg eticlopride.

(E) Comparison on D1 and D2 blockade (SCH23390, 0.1 mg/kg and eticlorpide, 0.16 mg/kg, respectively) after 12 days of initial training.

(F) Effect of initial training length on subsequent D2 blockade (eticlopride, 0.16 mg/kg) and recovery showing 12 (same data as E), 6, and 3 days of training. Throughout, red traces indicate saline controls. A 72 hr break occurred between treatment and drug-free recovery phases. n = 4/dose for (A–D) and n = 6 for (E and F); statistics reported in text. Error bars, SEM.

a consequence of different initial training lengths (treatment phase, $F_{(1,16)} = 3.02$, $p = 0.1$; recovery phase, $F_{(1,16)} = 0.53$, $p = 0.47$). The impairment and slow relearning induced subsequent to D2 blockade during recovery is more apparent with longer training periods as a consequence of having achieved higher performance during initial training; however, training under D2 blockade impairs subsequent performance independent of prior training (recovery, training length, $F_{(1,16)} = 0.53$, $p = 0.47$; training length x session, $F_{(1,16)} = 0.06$, $p = 0.80$), suggesting that the putative aberrant learning that is induced is dependent on experience during the D2 blockade and not prior skill level.

Together, the data in Figures 2 and 3 show a dissociation between the immediate, performance degrading effects of dopamine blockade and the subsequent, drug-free performance. These effects on subsequent behavior reflect learning and, presumably, synaptic plasticity that occurs during training under the dopamine blockade. Blockade of D1 appears to primarily induce a performance decrement with immediate recovery upon cessation of drug. In contrast, D2 blockade appears to induce an aberrant learning process that results in persistent impairment and slowed recovery in the drug-free condition.

We hypothesized that these differential effects of D1 versus D2 blockade could arise as a function of their different effects on activity and plasticity within the direct and indirect corticostriatal pathways, respective. Specifically, D1 blockade would reduce striatonigral MSN activity in the direct "GO" pathway, associated with selection of the correct actions, thus impairing performance; however, this overall decrease in activity may also protect against aberrant learning by decreasing the probability of Hebbian plasticity in the first place. In contrast, D2 blockade enhances excitability of striatopallidal MSNs in the indirect, "NOGO" pathway, increasing inhibitory activity that impairs performance. In this case, however, the increased activity may increase the probability of Hebbian plasticity (for review, Wiecki and Frank, 2010; Beeler, 2011), thus inducing aberrant learning in the striatopallidal, NOGO pathway.

## Adenosine Antagonists Mitigate Aberrant Learning

Although D2 activation is believed to be critical for the induction of synaptic depression (LTD) in D2-expressing MSNs, activation of A2A facilitates potentiation at these synapses (LTP) and A2A blockade prevents this potentiation (Shen et al., 2008; Lovinger, 2010; Peterson et al., 2012). We thus tested the effectiveness of A2A antagonists in mitigating the deficits induced by dopamine blockade as described above, by administering adenosine receptor antagonists during dopamine blockade.

We coadministered an A2A-selective antagonist, SCH58261, with the dopamine antagonists cocktail during initial acquisition and tested recovery under drug-free conditions. SCH58261 did not rescue performance during cocktail administration (Figure 4A, left; during acquisition, dose, $F_{(1,21)} = 0.025$, $p = 0.87$, dose X session, $F_{(1,21)} = 1.34$, $p = 0.25$). Performance during drug-free recovery, however, differed according to the dose of SCH58261 coadministered, though only a trend (Figures 4A

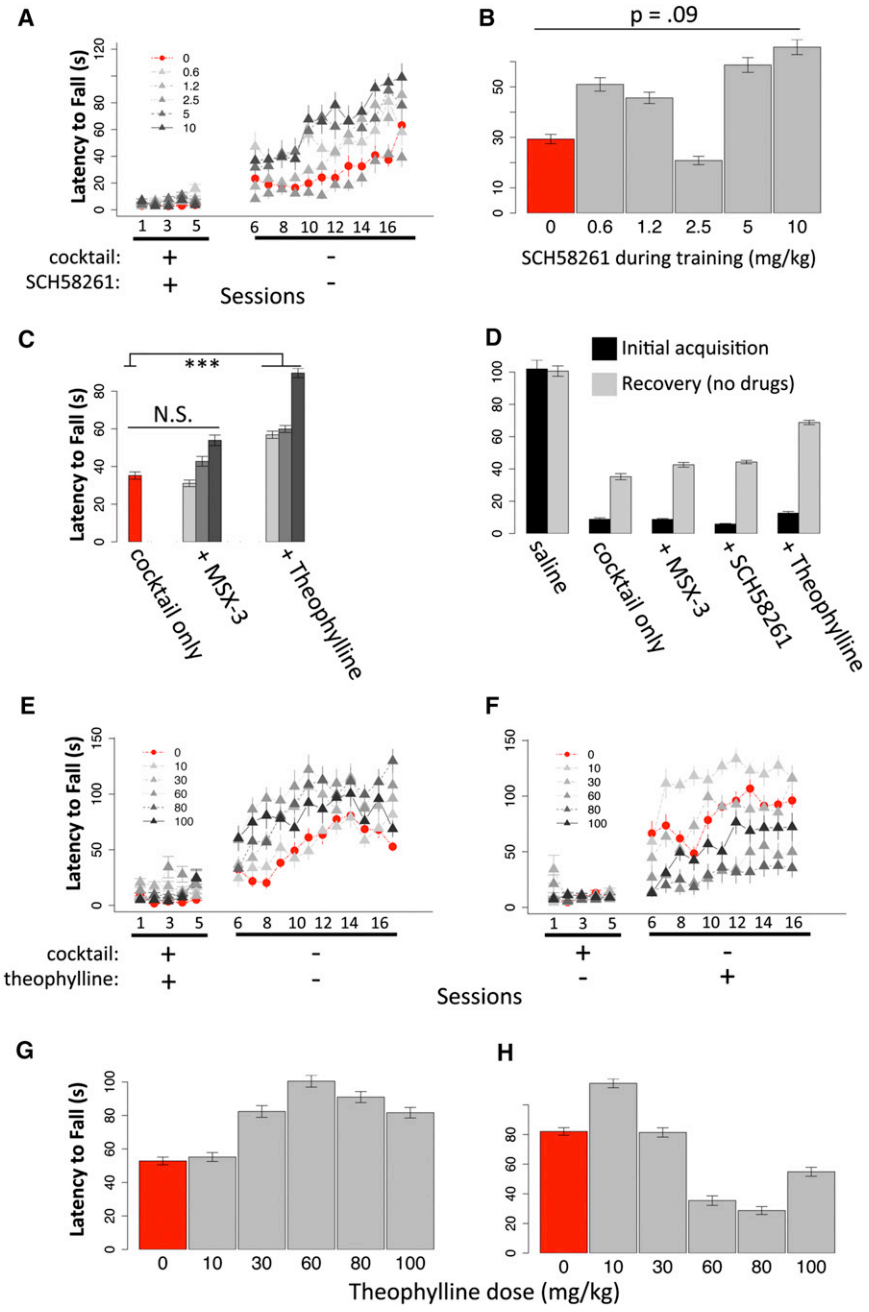


**Figure 4. Effect of Adenosine Antagonists on Impairment and Recovery from Dopamine Antagonist Cocktail Administered during Initial Acquisition**

(A) Latency to fall across consecutive days/sessions for mice coadministered SCH 58261 during cocktail training (red trace, cocktail only; gray darkens with increasing SCH58261) during the initial drug treatment phase (TX) and the drug-free recovery phase (NO TX), with a 72 hr break between phases.

(B) Summary of dose-dependent effects of SCH58261 coadministered during initial acquisition on subsequent drug-free recovery. Mean latency to fall averaged across sessions during the drug-free recovery phase plotted by SCH58261 dose (red bar, cocktail only).

(C) Mean latency to fall averaged across all drug-free recovery sessions showing dose response for MSX-3 and theophylline coadministered with dopamine antagonists cocktail during initial acquisition (red bar, mice administered cocktail only; darker gray shades represent increasing doses. MSX-3, 1, 2.5, and 5 mg/kg; SCH58261, 0.6, 1.2, 2.5, 5, and 10 mg/kg; theophylline, 10, 30, and 60 mg/kg).

(D) Summary of the average latency to fall averaged across doses and sessions during initial acquisition (black bars) and the subsequent drug-free recovery (gray bars) for mice coadministered MSX-3, SCH59261, or theophylline with a cocktail of SCH23390 (0.1 mg/kg) and eticlopride (0.16 mg/kg) during initial acquisition.

(E and F) Latency to fall in wild-type C57BL/6 mice administered the adenosine antagonist theophylline at the specified dose either (E) during initial acquisition under a cocktail of dopamine antagonists or (F) during the recovery phase subsequent to initial training under cocktail. Red traces represent control mice receiving no theophylline. Each point represents the average of five trials during daily sessions with a 72 hr break between phases.

(G and H) Mean performance across sessions during the recovery phase plotted by theophylline dose for mice administered theophylline during either (G) initial training under cocktail or (H) the recovery phase. n = 4/dose, statistics reported in text. ***p < 0.001. Error bars, SEM.

and 4B; dose, $F_{(1,21)} = 3.01$, $p = 0.09$; dose x session, $F_{(1,21)} = 1.28$, $p = 0.27$). It is striking that coadministered SCH58261, despite lack of observable effects during its administration, appears to improve subsequent performance during the drug-free recovery phase; however, the response does not appear to be linearly related to dose (Figure 4B).

We next screened two additional drugs, MSX-3, another selective A2A antagonist and theophylline, a nonselective adenosine receptor antagonist. Both drugs coadministered during training under dopamine antagonists cocktail appear to improve subsequent drug-free recovery (Figure 4C), though only theophylline is significantly different from cocktail alone (MSX-3 dose, $F_{(1,14)} = 1.65$, $p = 0.21$; theophylline dose, $F_{(1,14)} = 19.65$, $p < 0.001$).

In the remainder of the studies, we use the nonspecific adenosine antagonist theophylline over SCH58261 for several reasons. First, although not selective between A1 and A2A, numerous studies have demonstrated that motor-enhancing effects of nonselective adenosine antagonists are mediated through A2A and not A1 (El Yacoubi et al., 2000; Kelsey et al., 2009; Hsu et al., 2010). Moreover, in a study using 6-OHDA lesion animal models, both the nonspecific adenosine antagonist

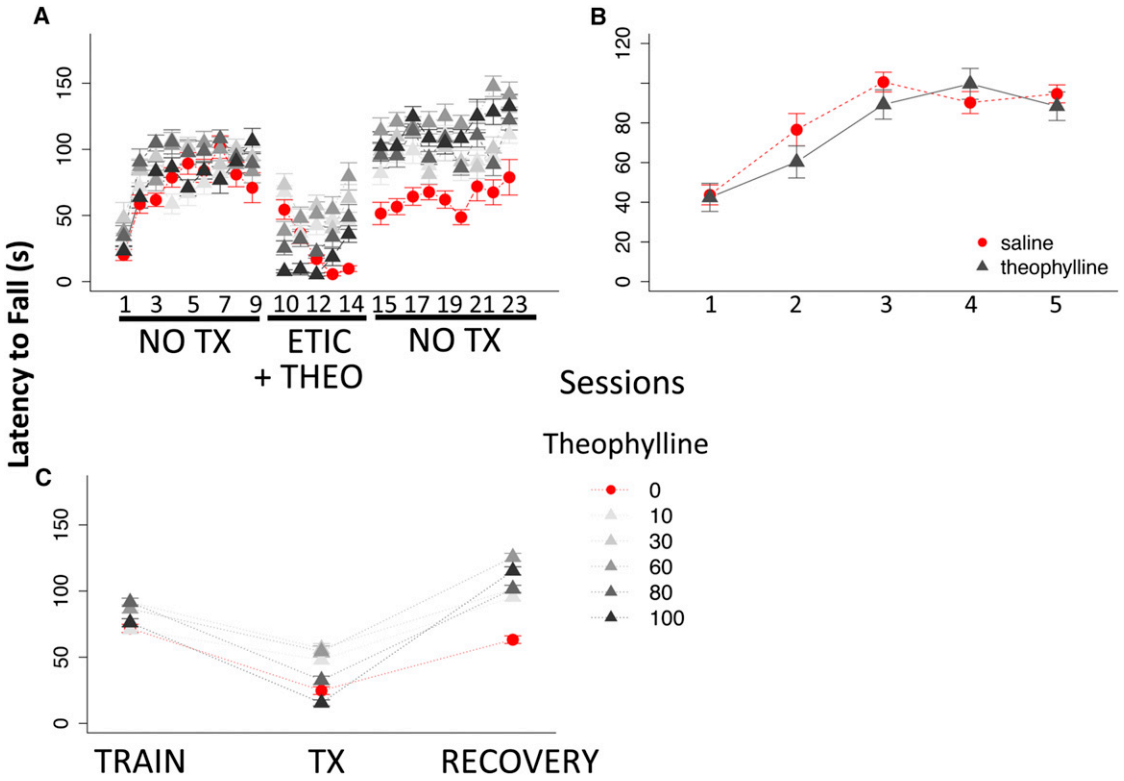


**Figure 5. Effect of Adenosine Antagonist Theophylline on Eticlopride-Induced Aberrant Learning and Recovery of an Established Skill**

(A) Latency to fall in wild-type C57BL/6 mice coadministered the A2A antagonist theophylline at the specified dose together with the D2 antagonist eticlopride (0.16 mg/kg) after initial drug-free training to asymptotic performance (i.e., established skill).

(B) Latency to fall in mice administered either theophylline at 80 mg/kg or saline for 5 days initial acquisition.

(C) Mean latency to fall across sessions during the different phases of the experiment. n = 4/dose, statistics reported in text. Error bars, SEM.

caffeine and the A2A-selective SCH58261 enhanced motor function and L-DOPA efficacy following the lesion, whereas the A1-selective antagonist CPT did not, again demonstrating that it is the A2A-specific actions of caffeine (and by extension, theophylline) that are relevant (Kelsey et al., 2009). Second, theophylline is a compound that has been in clinical use for decades, and there is clinical evidence that theophylline has therapeutic efficacy in PD (Mally and Stone, 1996; Kostic et al., 1999; Kulisevsky et al., 2002), making it more relevant to potential clinical studies. Finally, theophylline is water-soluble and does not require DMSO vehicle. In some cases, we administer treatments over a period of weeks and wanted to avoid the potential confound of chronic injections of DMSO. Thus, we chose theophylline to investigate the amelioration of aberrant learning.

### Theophylline Diminishes Aberrant Learning during Dopamine Blockade but Impairs Performance when Administered during Recovery

During initial acquisition, theophylline had little effect on the performance impairment induced by the D1/D2 antagonist cocktail (Figure 4E, left; dose main effect, $F_{(1,22)} = 0.90$, $p = 0.35$). However, during the drug-free recovery phase (neither cocktail nor theophylline administered), a dose-dependent improvement in recovery was observed (Figures 4E and 4G; dose main effect, $F_{(1,22)} = 3.4$, $p < 0.07$) similar to that observed with SCH58261 (Figure 4A), suggesting that theophylline diminished the putative aberrant learning that occurred during acquisition under conditions of dopamine blockade. In contrast, theophylline administered during the recovery phase, with the exception of the lowest dose, impairs recovery (Figure 4F; dose main effect, $F_{(1,22)} = 10.4$, $p < 0.01$), showing a dose-response curve that reflects almost the inverse mirror (Figure 4H) of that observed when theophylline is administered during acquisition (Figure 4G).

### Theophylline Protects Established Skills

We then tested whether theophylline modified the effects of dopamine receptor blockade on established motor skills by first training mice to asymptotic performance under normal conditions (i.e., no dopamine manipulation). Then eticlopride was administered to induce degradation in performance and putative aberrant learning (as in Figures 3B, 3E, and 3F), either with or without coadministration of theophylline. In mice that received only eticlopride, we observe the same gradual deterioration in performance and gradual recovery observed above (Figure 5A). In contrast, in mice that received coadministered theophylline, the performance impairment induced by D2 blockade did not show gradual deterioration but rather an immediate decrement

that remained constant across sessions (Figure 5A; during treatment, all groups, session x dose, $F_{(1,22)} = 13.1$, $p < 0.01$; each group separately, session was only significant for eticlopride-only group, $F_{(1,4)} = 12.33$, $p = 0.039$, all theophylline doses, session N.S.). Critically, upon discontinuation of the eticlopride and theophylline, the mice that had received theophylline showed no subsequent impairment in performance (Figures 5A and 5C; dose main effect on recovery, all groups, $F_{(1,22)} = 8.18$, $p < 0.01$; theophylline treated only, $F_{(1,17)} = 1.58$, $p = 0.227$), indicating that theophylline, though not eliminating the direct performance impairment induced by eticlopride, did effectively block aberrant learning protecting established skills.

### Adenosine Antagonism Has No Effect on Initial Acquisition under Normal Conditions

When applied during initial acquisition without dopamine blockade, theophylline has no effect on learning and performance (Figure 5B). These data suggest that a potential reduction in striatopallidal LTP arising from A2A blockade does not impair initial acquisition. Learning in the striatonigral pathway may compensate for reduced LTP in the indirect pathway in de novo learning; however, once aberrant learning is established, a full range of plasticity is apparently required to "unlearn" it and implement appropriate learning.

### D2 Blockade Induces Potentiation in Striatopallidal MSNs that Is Diminished by Theophylline

Corticostriatal LTD, believed to be critical for behavior and motor execution, is facilitated by activation of D2 receptors (Gerdeman et al., 2002; Calabresi et al., 2007; Kreitzer and Malenka, 2008; Shen et al., 2008; Lovinger, 2010). Loss of D2 activation, such as occurs with dopamine denervation or depletion, blocks high-frequency stimulation-induced LTD (Calabresi et al., 1997; Gerdeman et al., 2002; Kreitzer and Malenka, 2007; Bagetta et al., 2011). However, increasing evidence suggests that dopamine denervation not only impairs LTD but inverts it such that conditions that would normally induce LTD (i.e., high-frequency stimulation or spike-timing dependent LTD) instead induce LTP (Calabresi et al., 1997; Picconi et al., 2003; Shen et al., 2008; Peterson et al., 2012). These studies provide compelling evidence that loss of dopamine signaling at D2 receptors can invert corticostriatal plasticity in the striatopalidal pathway, which we propose underlies aberrant learning. Therefore, we tested the effects of D2 blockade on the strength of excitatory cortical inputs to stratopallidal MSNs.

In the presence of bath-applied sulpride (20 μM), we observed a gradual increase in the strength of glutamatergic input to D2-expressing MSNs, without any stimulation protocol (Figures 6B and 6C). As the measurement of EPSCs itself constitutes a low-frequency stimulation, for a subset of cells we withheld stimulation for 10 min during the sulpiride administration and applied two doses of sulpiride, 2 and 20 μM. In this subset of cells, we observe dose-dependent potentiation in the absence of any exogenously applied stimulation (Figure 6D). Previous reports demonstrate that LTP in the striatopallidal pathway is facilitated by activation of the A2A adenosine receptor and can be blocked by A2A antagonism (Schiffmann et al., 2003; Shen et al., 2008; Peterson et al., 2012). We then tested whether D2 blockade-induced potentiation is sensitive to A2A antagonism. Bath application of theophylline (1 μM) in combination with sulpiride reduced the potentiation seen with sulpiride alone (Figures 6B and 6C).

These data suggest that D2 blockade induces potentiation in striatopallidal MSNs, a process that may contribute to the increased responsiveness observed in indirect pathway MSNs after dopamine denervation (Mallet et al., 2005; Gertler et al., 2008; Peterson et al., 2012). These data are consistent with previous reports showing that dopamine denervation or loss of D2 signaling can induce LTP instead of LTD (Calabresi et al., 1997; Picconi et al., 2003; Shen et al., 2008; Peterson et al., 2012), suggesting that dopamine denervation and subsequent decreases in D2 signaling favor LTP in the corticostriatal synapses in the indirect pathway and increase inhibitory tone-modulating cortical activity. Enhanced excitatory drive onto striatopalidal MSNs under these conditions is consistent with the proposed mechanism underlying the aberrant learning observed behaviorally. The decrease in D2 blockade-induced potentiation observed with application of theophylline, together with reports that A2A antagonism can impede striatopallidal potentiation (Schiffmann et al., 2003; Shen et al., 2008; Peterson et al., 2012), is a likely explanation for the partial protection from aberrant learning observed in our behavioral studies.

### Modeling Rotarod Skill Learning

Adapting a previously published basal ganglia (BG) model (Frank, 2005; Wiecki et al., 2009), we simulated the demands of the rotarod task by assuming that the mice have to select between four possible motor outputs (R1–R4, e.g., which paw to move forward), depending on the sensory state. Motor tasks like the rotarod are dynamic and integrative such that the correct action needed in response to a particular sensory state (e.g., position on the rod) may depend on the identity of another variable (e.g., proprioceptive input). We simulated this type of task by including two sets of inputs, each having two possible stimuli (i.e., $S_{A1}$, $S_{A2}$; $S_{B1}$, $S_{B2}$). The "correct" motor actions (i.e., those that would prevent the animal from falling off the rod) were dependent on conjunctive combinations of the two sets. For example, if $S_{A1}$ is present, then $S_{B1}$ should be associated with R1 and $S_{B2}$ with R2. If $S_{A2}$ is present, then $S_{B1}$ should be associated with R3 and $S_{B2}$ with R4. We adopted this input-output structure to capture the integrative stimulus-response learning attributed to the dorsal striatum. The different sets of inputs to be integrated could potentially represent context and discrete stimuli, information from different sensory modalities (visual, proprioceptive, vestibular), different coordinate systems (e.g., position on rod as medial/lateral and forward/backward) and so on. The requirement to integrate on-going stimuli to determine a complex state and the appropriate response more realistically reflects the sensorimotor integration required in the rotarod task. To simulate the acceleration of the rotarod, we restricted the amount of time that the model had to select a response such that with each correct action the time limit was decreased. The model is described briefly in Figure 7 (schematic), and detailed description and equations are provided in the Extended Experimental Procedures.

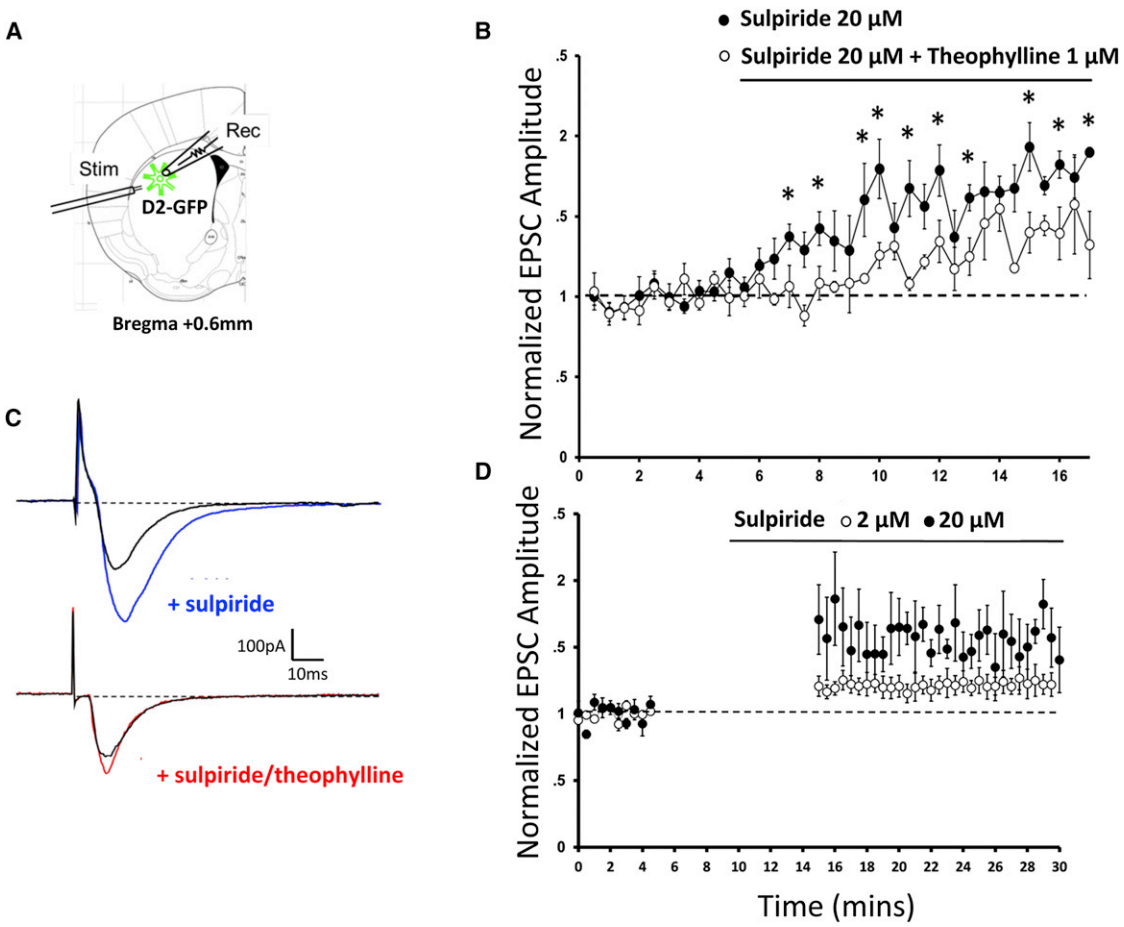


**Figure 6. D2 Blockade Potentiates Excitatory Inputs to Striatopallidal MSNs, which Is Reduced by Theophylline**

(A) D2-GFP medium spiny neurons in the dorsolateral striatum were held in voltage clamp (Vm, −70 mV; Rec, recording electrode). Stimulating electrodes (Stim) were placed near the corpus callosum, which allowed stimulation of corticostriatal evoked excitatory synaptic currents (EPSCs) at 30 s intervals.

(B) After baseline EPSC amplitude was established, bath application of the D2 receptor antagonist sulpiride potentiated evoked EPSC's (20 μM; filled symbols; n = 4). Coapplication of the adenosine antagonist theophylline (1 μM) significantly attenuated the effects of sulpiride on EPSC amplitude (open symbols; n = 6).

(C) Example traces from each recording condition.

(D) In a separate set of experiments, bath application of sulpiride still potentiated evoked EPSC's in the absence of any stimulation (solid symbols, n = 5). Similar effects were observed with administration of 2 μM sulpiride (open symbols, n = 3). *p < 0.05. Error bars, SEM.

## Modeling Dopamine Effects on Performance and Learning

Dopamine modulates the balance of activity and plasticity in the direct (GO) and indirect (NOGO) pathways. Increased dopamine excites the D1-expressing GO pathway and inhibits the D2-expressing NOGO pathway, whereas decreased dopamine has the opposite effect. Dopamine also modulates learning. Following a correct response, phasic increases modulate plasticity differently in the direct, striatonigral GO and the indirect, striatopallidal NOGO pathways. Increased phasic dopamine activates D1 and D2, increasing (LTP) and decreasing (LTD) synaptic weights in the direct and indirect pathways, respectively, in proportion to their activity, thereby facilitating (GO LTP) and disinhibiting (NOGO LTD) that response in future presentations of the same stimuli. The net effect is to drive activity-dependent plasticity so that weights from the active inputs to GO units associated with rewarding actions are increased, while diminishing the NOGO activity associated with those same inputs. When the network selects an erroneous response, phasic dips of dopamine induce the reverse learning process such that the weights in the indirect, striatopallidal NOGO pathway are strengthened (LTP), reducing the likelihood of repeating the error in the future.

These performance and learning effects are interactive. For example, D2 blockade enhances the excitability of NOGO units and hence their propensity for activity-dependent plasticity. This effect manifests itself such that even if adaptive actions are selected, the greater NOGO activity arising from D2 blockade drives inhibitory learning in response to the current sensory states but does so as if there had been a dip in phasic dopamine activity. Thus, even after D2 blockade is removed, NOGO activity is associated with correct responses and impairs performance.

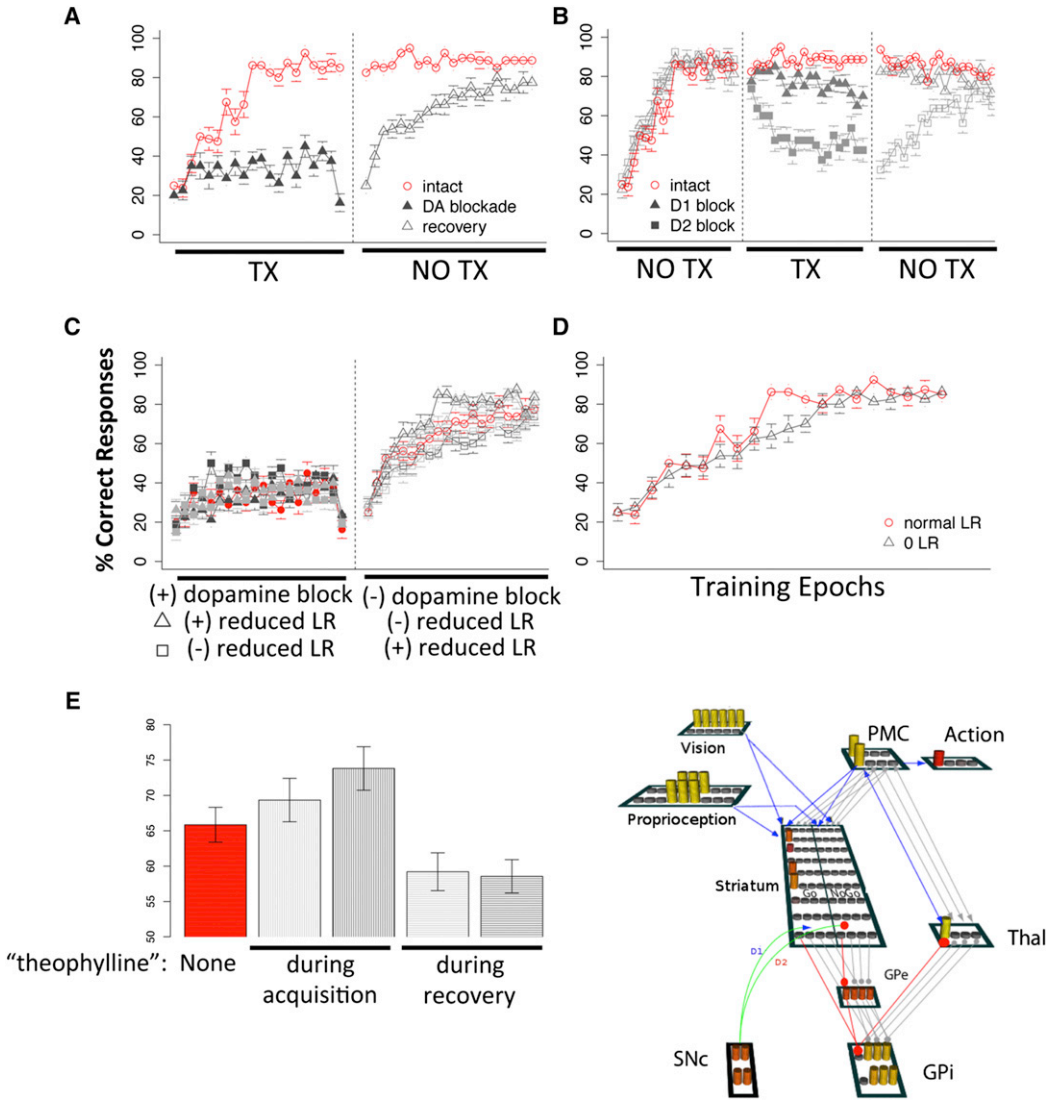


**Figure 7. Model Performance under Conditions Recapitulating Mouse Experiments**

(A) Shows the percentage of correct responses by the model intact (open red symbols) and with dopamine blockade (filled gray triangles) and subsequent recovery with dopamine activity restored (open gray triangles). Each point represents the average of four trials.

(B) Shows the effect of blocking either the D1/GO layer (filled triangles) or the D2/NOGO layer (filled squares) subsequent to initial intact learning and subsequent recovery when dopamine function is restored (open symbols).

(C) The model was trained under dopamine blockade with a reduced learning rate in the D2/NOGO layer (learning set to one-half, light gray, set to zero, dark gray) to simulate A2A antagonism, either during initial acquisition under dopamine blockade (triangles) or during dopamine-restored recovery (squares).

(D) Initial model learning and performance under normal dopamine function with (gray symbols) and without (red symbols) D2/NOGO learning rate reduced to zero to simulate A2A antagonism.

(E) Bar graph showing average performance across epochs during the recovery phase grouped by time (during acquisition or recovery) and degree of reduction in D2/NOGO learning rate (i.e., "theophylline").

Schematic of basal ganglia neurocomputational model. The basal ganglia model includes layers incorporating the direct (GO) and indirect (NOGO) pathways from cortex (two input layers) through the striatum (GO and NOGO units), to the globus pallidus externa (GPe), the substantia nigra reticulata/globus pallidus interna (GPi), the thalamus, the premotor cortex (PMC) to the output, or motor cortex. SNc dopamine neurons project to both the GO and NOGO layers of the striatum simulating projections to D1 and D2 MSNs (GO/NOGO, respectively, see the Extended Experimental Procedures for a detailed description). Fast-spiking inhibitory interneurons (data not shown) regulate activity in both striatal populations via feed-forward inhibition. At each trial, the network is presented input from each of two input layers (raised cylinders represent example unit activity). Premotor cortical (PMC) units representing the four candidate responses then become noisily activated. Under baseline conditions, the thalamus is inhibited, but a response is selected once a thalamic unit becomes disinhibited and amplifies activity in the corresponding motor units. The role of the BG is to modulate activity in the thalamus according to whether the responses are adaptive in the current sensory state. Each phase of the experiment (i.e., analogous to with or without administered drugs, as in the mouse studies) consisted of 20 epochs, equivalent to sessions. Each epoch contained four trials. Model performance is reported as a percentage of correct responses. Each data point is the average performance of 20 models initialized with different random weights (i.e., n = 20). Error bars, SEM.

### Simulating Pharmacological Manipulations

Pharmacological blockade of D1 or D2 were directly and independently mimicked in the model by reducing the efficacy with which dopamine modulates the excitatory D1-GO projections or the inhibitory D2-NOGO projections to approximately 65% of the intact values (Extended Experimental Procedures). As the partial protection against aberrant learning observed above is hypothesized to arise from an A2A blockade that diminishes plasticity in the indirect, D2-expressing pathway, we examined whether the effects of theophylline could be recapitulated solely by reducing or eliminating plasticity in the corticostriatal projections to D2 units. The model then tests whether plasticity in the indirect pathway alone can account for the behavioral findings.

### Model Recapitulates Effects of Dopamine Blockade on Learning and Performance

Figure 7A shows that the intact model can robustly learn the conjunctive associations and select correct actions R1–R4. As expected, simulated D1/D2 blockade results in the same severe performance deficit observed in the mouse studies (Figure 7A, left). However, we hypothesized that it would also drive aberrant learning as described above. Indeed, upon removal of the blockade, recovery is gradual (slower than that for intact models during initial acquisition) and does not fully recover (Figure 7A, middle).

We next examined whether the model can account for performance if dopamine blockade is applied after asymptotic performance, independently for D1 and D2 receptors (Figure 7B). We observed that initial application of D2 blockade after learning was associated with relatively preserved motor performance that then gradually declined, as observed with mice. With D1 blockade, there is less impairment and, unlike in the mice, it does show a small gradual decline, though not as pronounced as D2 blockade. However, upon removal of D1 blockade, performance returned immediately to prior asymptotic levels, whereas with D2 blockade a gradual relearning was required, consistent with behavioral observations.

The model recapitulates the pharmacological data because D2 blockade enhances activity in the NOGO units and induces synaptic potentiation (LTP) even following correct responses that would normally induce depression (LTD), driving aberrant learning. Once indirect pathway inhibition has been learned through inappropriate LTP, it needs to be unlearned for recovery to occur even when dopamine is restored. In contrast, D1 blockade, although it prevents the network from selecting the correct response due to reduced facilitation from GO units, does not significantly induce an aberrant learning process because D1 blockade reduces activity and lowers the propensity for activity-dependent plasticity, thus minimizing aberrant learning. Thus, in contrast to D2 blockade, removal of D1 blockade allows the network to express its previously learned adaptive weights.

### Can Blocking Plasticity in the D2 Pathway Mimic Theophylline Administration?

Finally, we tested whether plasticity in the indirect, NOGO pathway can account for the differential effects of theophylline when the drug is applied during acquisition or recovery. Specifically, as A2A antagonism can diminish striatopallidal LTP under at least some conditions (Schiffmann et al., 2003; Shen et al., 2008; Peterson et al., 2012), we modeled theophylline as a reduction of plasticity in the striatopallidal NOGO units.

Consistent with behavioral observations, simulated A2A antagonism (blocking corticostriatal NOGO plasticity) during initial acquisition did not rescue performance during simultaneous dopamine blockade (Figure 7C, left). However, blockade of NOGO plasticity did improve subsequent recovery when dopamine blockade was removed, reducing aberrant learning (Figures 7C and 7E). Moreover, subsequent recovery speed was now in the same range as that of intact models during initial learning (compare with previous Figure 7A) and better than recovery without NOGO blockade of plasticity (Figure 7C).

In the mouse studies, theophylline applied during the recovery phase rather than during acquisition impaired rather than improved recovery. The same effect was observed in the model (Figures 7C and 7E). Reducing or blocking plasticity in the NOGO pathway during recovery impairs the model's ability to (1) learn which actions should be suppressed because they are maladaptive and (2) "unlearn" inhibitory, NOGO weights inappropriately associated with correct actions. Finally, as with the mice, blocking plasticity in the NOGO pathway had little effect during initial acquisition under normal conditions (i.e., no dopamine blockade, Figure 7D), suggesting in the naive state, correct stimulus-action associations can be learned in the GO pathway, suggesting that aberrant learning and plasticity in the indirect pathway (i.e., inappropriate LTP) is more deleterious than a simple deficit of plasticity.

Unlike experiments with animals and humans, where it is difficult to isolate potential mechanisms and substrates, in a computational model we have complete control. Using an a priori model of the basal ganglia that has been applied to various human and animal data sets, we show here that the aberrant learning observed in mice can be accounted for specifically by alterations in plasticity in the indirect, striatopallidal pathway. As this model was neither developed specifically for these studies nor to test the aberrant learning hypothesis, it is notable that the model recapitulates so closely the array of behavioral phenomena observed.

## DISCUSSION

Corticostiatal throughput is modulated through two main pathways. Activity in the D1-expressing direct, striatonigral "GO" pathway favors disinhibition of cortical activity and facilitates behavioral throughput. Activity in the D2-expressing, indirect striatopallidal "NOGO" pathway, in contrast, favors inhibition of cortical activity and inhibits behavioral throughput (Albin et al., 1989; Alexander et al., 1990; Mink, 1996). Dopamine shifts the balance between these two pathways such that increased dopamine increases the responsiveness of the GO pathway via D1 activation, while simultaneously decreasing the influence of the NOGO pathway via activation of D2. Conversely, diminished dopamine will favor the inhibitory NOGO pathway because of greater activity in D2-expressing MSNs as a consequence of less activation of D2. Much evidence supports this dual pathway

model in controlling motor performance in rodents (Hikida et al., 2010; Kravitz et al., 2010).

Dopamine also modulates corticostiatal plasticity (Calabresi et al., 1992a, 1992b; Kreitzer and Malenka, 2008; Surmeier et al., 2009; Wickens, 2009; Lovinger, 2010) in both the striatonigral GO and striatopallidal NOGO pathways, further influencing motor performance through learning. Corticostriatal plasticity can enhance or diminish the responsiveness of either pathway to cortical input, selectively facilitating the expression or inhibition of specific responses and motor skills. Our previous studies suggest that this adaptive plasticity is altered in the D2 pathway under conditions of dopamine depletion, blockade, or denervation (Wiecki et al., 2009; Beeler et al., 2010; Wiecki and Frank, 2010; Beeler, 2011), giving rise to an aberrant learning that selectively encodes inappropriate inhibition through experience-dependent synaptic changes that impede future motor responses even if dopamine is restored (i.e., such as in L-DOPA treatment).

We propose that the mechanism that underlies aberrant learning is altered corticostriatal plasticity in the indirect, striatopallidal pathway that favors synaptic potentiation (LTP) at the expense of synaptic depression (LTD), inappropriately increasing the responsiveness of striatopallidal MSNs to cortical input and pathologically increasing behavioral inhibition. Substantial evidence supports this proposed mechanism. In mice lacking D2 receptors, the same high-frequency stimulation protocol (HFS) that induces LTD instead induces LTP (Calabresi et al., 1997). In 6-OHDA lesioned rats, a model of PD, HFS also induces LTP rather than the normal LTD (Picconi et al., 2003), though Kreitzer and Malenka (Kreitzer and Malenka, 2007) in a similar study observed only a loss of LTD and not its inversion to LTP. The reason for this discrepancy is unclear, though subsequent reports confirm the inversion of plasticity under dopamine denervation and further suggest the abnormal LTP is dependent upon A2A activation (Shen et al., 2008; Peterson et al., 2012). Here, we show that in the absence of HFS, D2 blockade induces potentiation in striatopallidal synapses that is diminished by administration of an adenosine antagonist that blocks A2A, consistent with recent studies and with the observed ameliorative effects of theophylline administered during dopamine blockade. Taken together, the present data and published studies strongly support the hypothesis that D2 blockade (and dopamine denervation/depletion) shift striatopallidal plasticity to inappropriately favor LTP. By favoring and increasing synaptic potentiation at the expense of depression, dopamine blockade/denervation increases the learned responsiveness of striatopallidal neurons to afferent input, consequently enhancing inhibitory tone on cortical activity, as proposed to underlie aberrant learning (Wiecki et al., 2009; Beeler, 2011). In short, dopamine denervation is widely believed to induce an imbalance between the direct and indirect pathways; here, we suggest that altered corticostriatal plasticity contributes a learned component to this imbalance: inappropriate inhibition structurally embedded as physical changes in synapses. Importantly, though pharmacological treatment may both reverse the imbalanace in activity between the direct and indirect pathways and restore normal plasticity, structural changes arising from aberrant learning can only be reversed through further structural changes, that is, relearning.

We attribute the partial protection against aberrant learning afforded by theophylline to its actions on postsynaptic A2A receptors. Though theophylline is a nonselective adenosine antagonist, several studies have demonstrated that the motor-improving effects of nonselective adenosine antagonists are mediated through their actions on A2A; A2A- but not A1-selective compounds yield the same results (El Yacoubi et al., 2000; Kelsey et al., 2009; Hsu et al., 2010). Moreover, published studies have demonstrated that A2A blockade can decrease potentiation in striatopallidal MSNs (Shen et al., 2008; Peterson et al., 2012), consistent with the electrophysiological data obtained here. As A2A is expressed both postsynaptically on striatopallidal MSNs, as well as presynaptically on glutamatergic terminals, it is possible that the effects of theophylline, even on A2A receptors, is mediated pre- rather than postsynaptically. Indeed, Calabresi and colleagues have demonstrated that A2A antagonist (though notably only in combination with D2 agonist) can reduce glutamatergic transmission through a presynaptic mechanism (Tozzi et al., 2007); however, Quiroz et al. (2009) have shown that the A2A receptor is expressed presynaptically only on cortical afferents synapsing on striatonigral MSNs, suggesting that effects of A2A antagonist on striatopallidal cells will be mediated exclusively through postsynaptic A2A receptors. Finally, the model suggests that reducing synaptic plasticity in the striatopallidal pathway can account for and recapitulate the theophylline effects, further supporting a striatopallidal, postsynaptic mechanism underlying the partial protection from aberrant learning conferred by theophylline. Nonetheless, it should be borne in mind that theophylline is not A2A selective, leaving open the possibility that other actions in addition to A2A antagonism may contribute to its ameliorative effects.

The pharmacological model used here captures one aspect of PD, the reduction in activation of D1 and D2 that occurs as a consequence of denervation. The simplicity of the model, the reversibility of pharmacological blockade together with ability to block D1 and D2 individually represents strengths that facilitated the present studies, which would have been intractable in the more traditional 6-OHDA model of PD. We show that impaired dopamine signaling induces aberrant learning. By necessity, our model induces acute, temporary decreases in dopamine signaling. The role such aberrant learning plays in PD is likely to be complex and complicated by adaptation to chronic denervation; however, in a previous study using PITx3-deficient mice that exhibit a 90% loss of dopamine denervation in the dorsal striatum from birth and show physiological adaptations characteristic of PD, we observe a similar aberrant learning (Beeler et al., 2010). We have advanced the hypothesis that the poorly understood long-duration response (LDR) to L-DOPA (Muenter and Tyce, 1971; Anderson and Nutt, 2011) arises as a correction of aberrant learning (Beeler et al., 2010; Beeler, 2011). In this view, under progressive dopamine denervation, patients take a ''double hit'' in that declining dopamine induces direct motor performance deterioration but also through abnormal corticostriatal plasticity and aberrant learning, unravels previously established learning, essentially inverting it to favor inhibition rather than facilitation of movement. The short-duration response (SDR) of L-DOPA then reflects the

correction of direct motor performance deficits induced by dopamine depletion and lasts only as long as L-DOPA is present (Nutt et al., 1997). In contrast, by correcting underlying abnormal corticostriatal plasticity, L-DOPA restores normal learning and skill building. This corrective aspect of L-DOPA treatment is cumulative and retained as the appropriate calibration of millions of synaptic strengths that endures during trough periods of medication—or even on discontinuation of treatment, until aberrant learning reverses this learning and inappropriate synaptic strengths again predominate.

This view suggests an alternative therapeutic strategy of targeting pathways that mediate abnormal corticostriatal plasticity in the D2-expressing, indirect pathway; in essence, seeking to effect an LDR-like therapeutic independent of an SDR-like effect. The theophylline studies described here suggest this strategy may be feasible. Theophylline improved recovery in a dose-dependent manner when administered during putative aberrant learning under dopamine blockade, suggesting this intervention partially mitigated aberrant learning. In contrast, when administered subsequent to aberrant learning, it did not facilitate performance; indeed, at most doses it appeared to slow recovery. These data would suggest that A2A antagonists are likely to have limited therapeutic efficacy in ameliorating the direct effects of dopamine denervation on performance, that is, limited SDR-like actions. However, A2A antagonists may mitigate underlying abnormal corticostriatal plasticity in the D2-expressing indirect pathway and diminish aberrant learning, as observed here, inducing an LDR-like therapeutic efficacy. Importantly, from a clinical perspective, the A2A antagonism had no effect on learning under normal conditions (i.e., without prior aberrant learning). These observations may suggest alternative perspectives on the clinical potential of A2A antagonists, including their use early in the disease process to preserve established skills and slow aberrant learning, though more investigation is necessary.

The present observations have implications for rehabilitative approaches to treating PD (Abbruzzese et al., 2009; Keus et al., 2009; Nieuwboer et al., 2009). Rehabilitation protocols are, at their core, based on repetition and practice. If dopamine depletion induces experience-dependent aberrant learning, then it is possible that skill practice during low medication states (e.g., medication troughs, drug holidays) might degrade rather than improve motor skills. In contrast, skill practice during peak medication, when normal corticostriatal plasticity and learning are restored, would facilitate optimal performance gains. In short, rehabilitative treatments may potentially enhance or diminish the LDR of L-DOPA treatment, depending upon when they are administered. The "use it or lose it" strategy (Archer et al., 2011) may depend critically on timing.

Evidence suggests that aberrant learning may precede frank motor symptoms of PD (see Beeler, 2011 for review). Because different territories of the striatum are differentially affected as denervation progresses, it is possible that a process of denervation → aberrant learning → compensation → failure of compensation is recapitulated across striatal territories at different time courses. Thus, aberrant learning may play an important role in the development of cognitive impairments in PD (Wiecki and Frank, 2010). Consequently, an "LDR-like" treatment targeting aberrant learning may be useful in ameliorating or delaying the development of cognitive symptoms.

## EXPERIMENTAL PROCEDURES

### Animals

Mice were housed in standard conditions on a 06:00 to 18:00 light cycle with ad libitum food and water. Experiments were carried out during the light cycle. Animal procedures were approved by the Institutional Animal Care and Use Committee at the University of Chicago. All mice were C57BL/6 wild-type mice between 8–12 weeks of age. For the in vitro electrophysiology, adult (4- to 12-week-old) transgenic mice hemizygous for Drd2-enhanced green fluorescent protein (EGFP) bacterial artificial chromosome (BAC) of both sexes were used in all experiments. Drd2-EGFP homozygotes were identified from test crosses and then crossed with C57BLK6 mice to produce hemizygotes.

### Behavior Tests

A computer-controlled rotarod apparatus (Rotamex-5, Columbus Instruments, Columbus, OH, USA) with a rat rod (7 cm diameter) was set to accelerate from 4 to 40 revolutions per minute rpm) over 300 s, and recorded time to fall. Mice received five consecutive trials per session, one session per day (~30 s between trials). Open field chambers were 40 × 40 cm (Med Associates, St. Albans, VT, USA) with lighting at 21 lux. Infrared beams recorded the animals' locomotor activity and rearing movements (vertical activity). Data were collected in 5 min bins during 45 min sessions for 3 days. For the treadmill performance, a plexiglass enclosure was placed over the treadmill to force the mice to remain on the track during the trials (enclosed treadmill space, 5 × 20 cm). Mice were provided three 20 s trials per day for 3 days at incrementing speeds (10, 15, and 20; 15, 15, and 20; 15, 20, and 20 cm/s, days 1–3, respectively). Their performance was assessed on the final day at 15 and 20 cm/s by recording the amount of time the animals remained in the forward two-thirds of the track versus falling back into the back third.

### Drug Administration

All injections were intraperitoneal (i.p.) at 0.01 ml/g of body weight. SCH23390, eticlopride, and theophylline (Sigma-Aldrich, St. Louis, MO, USA) were administered 30 min prior to sessions at the specified doses prepared in 0.9% saline.

### In Vitro Electrophysiology

After isoflurane anesthesia, animals were decapitated, and their brains removed to ice-cold sucrose aCSF (in mM: sucrose 125, KCl 2.5, MgCl2 1, CaCl2 2.5, glucose 20, NaH2PO4 1, NaHCO3 25, ascorbic acid 10; bubbled with 95% O2/5% CO2). Coronal slices were cut (250 μM; VT1000S, Leica) containing the dorsolateral striatum (+0.4 mm-+1.0 mm from bregma) and removed to a holding chamber perfused with normal aCSF (in mM: NaCl 125, KCl 2.5, MgCl2 1, CaCl2 2.5, glucose 20, NaH2PO4 1, NaHCO3 25, ascorbic acid 1; bubbled with 95% O2/5% CO2), 20 ml/min at 34°C. For recording, the slices were superfused with normal aCSF without ascorbic acid at 2 ml/min, 32°C. In the dorsolateral striatum D2-GFP MSN's were visualized using fluorescence illumination on an upright microscope (Axioskop, Zeiss, Oberkochen, Germany). Whole-cell voltage clamp recordings used an Axopatch 200B amplifier, a Digidata 1200 interface, and pCLAMP 8 (Molecular Devices, Sunnyvale, CA, USA). All recordings were filtered at 1 kHz and digitized at 5kHz, Vm = −70 mV. For all recordings we used borosilicate electrodes (3–7 MΩ) containing (in mM), K-gluconate 154, KCl 1, EGTA 1, HEPES 10, glucose 10, and ATP 5 (pH 7.4 with KOH). Only cells with series resistance <20 MΩ were included. For stimulation of corticostriatal inputs, a bipolar tungsten electrode with a 500 μm tip separation was placed inside the cortical border of the dorsolateral striatum. After establishing a consistently evoked EPSC amplitude once every 30 s for 5 min, sulpiride (20 μM) was bath applied either on its own or with theophylline (1 μM). Recording was continued in the presence of drug for a further 15–20 min after which most of the recordings were terminated. Sulpiride and theophylline were purchased from Sigma-Aldrich.

### Statistical Analysis

#### *Behavior*

In all rotarod studies, the data were tested for significance using ANOVA (R statistical software [R version 2.12.1 2010-12-16] The R Foundation for Statistical Computing; http://www.r-project.org). Dose effects were modeled as continuous; comparison between drugs (including drug/dose comparisons) were modeled as factors with a level for either each drug or each drug/dose combination. As we used repeated-measures (five trials/session, multiple sessions across experiment), session was always included as a continuous independent variable with an error term of mouse/session for repeated-measures, for example,

$$\text{statistics} = \text{aov}(\text{latency} \sim \text{drug} * \text{session} + \text{Error}(\text{mouse}/\text{session})).$$

#### *Electrophysiology*

All data are reported as mean ± SEM and expressed as the normalized value of the baseline (5 min) before drug application. For analysis of drug effects on EPSC amplitude we used a two-way repeated-measures ANOVA, followed by a Tukey post hoc test on the EPSC amplitudes collected 5 min before drug exposure and 15–20 min after that time. Statistical significance was determined by $p < 0.05$; all statistical tests were performed using Sigmastat (Systat software).

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2012.11.014.

## LICENSING INFORMATION



## ACKNOWLEDGMENTS

This work was support by the National Institute of Drug Abuse (NIDA) (DA25875 to J.A.B.) and the National Institute of Neurological Disorders and Stroke (NINDS) (R21NS070269 to X.Z.). We would like to acknowledge the contribution of Zhen Fang, Huang Cao, Jessica Koranda, and Mari Murakami to developing the rotarod method and strategy and Carrie Swetlik for assistance with additional behavior testing. We thank Giselle Petzinger for insightful feedback on the manuscript.

Received: May 21, 2012
Revised: August 22, 2012
Accepted: November 19, 2012
Published: December 13, 2012

## REFERENCES

Abbruzzese, G., Trompetto, C., and Marinelli, L. (2009). The rationale for motor learning in Parkinson's disease. Eur. J. Phys. Rehabil. Med. *45*, 209–214.

Albin, R.L., Young, A.B., and Penney, J.B. (1989). The functional anatomy of basal ganglia disorders. Trends Neurosci. *12*, 366–375.

Alexander, G.E., Crutcher, M.D., and DeLong, M.R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. Prog. Brain Res. *85*, 119–146.

Anderson, E., and Nutt, J. (2011). The long-duration response to levodopa: phenomenology, potential mechanisms and clinical implications. Parkinsonism Relat. Disord. *17*, 587–592.

Archer, T., Fredriksson, A., and Johansson, B. (2011). Exercise alleviates Parkinsonism: clinical and laboratory evidence. Acta Neurol. Scand. *123*, 73–84.

Bagetta, V., Picconi, B., Marinucci, S., Sgobio, C., Pendolino, V., Ghiglieri, V., Fusco, F.R., Giampà, C., and Calabresi, P. (2011). Dopamine-dependent long-term depression is expressed in striatal spiny neurons of both direct and indirect pathways: implications for Parkinson's disease. J. Neurosci. *31*, 12513–12522.

Balleine, B.W., and O'Doherty, J.P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology *35*, 48–69.

Beeler, J.A. (2011). Preservation of function in Parkinson's disease: what's learning got to do with it? Brain Res. *1423*, 96–113.

Beeler, J.A., Cao, Z.F.H., Kheirbek, M.A., Ding, Y., Koranda, J., Murakami, M., Kang, U.J., and Zhuang, X. (2010). Dopamine-dependent motor learning: insight into levodopa's long-duration response. Ann. Neurol. *67*, 639–647.

Bernheimer, H., Birkmayer, W., Hornykiewicz, O., Jellinger, K., and Seitelberger, F. (1973). Brain dopamine and the syndromes of Parkinson and Huntington. Clinical, morphological and neurochemical correlations. J. Neurol. Sci. *20*, 415–455.

Bódi, N., Kéri, S., Nagy, H., Moustafa, A., Myers, C.E., Daw, N., Dibó, G., Takáts, A., Bereczki, D., and Gluck, M.A. (2009). Reward-learning and the novelty-seeking personality: a between- and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. Brain *132*, 2385–2395.

Calabresi, P., Maj, R., Pisani, A., Mercuri, N.B., and Bernardi, G. (1992a). Long-term synaptic depression in the striatum: physiological and pharmacological characterization. J. Neurosci. *12*, 4224–4233.

Calabresi, P., Pisani, A., Mercuri, N.B., and Bernardi, G. (1992b). Long-term Potentiation in the Striatum is Unmasked by Removing the Voltage-dependent Magnesium Block of NMDA Receptor Channels. Eur. J. Neurosci. *4*, 929–935.

Calabresi, P., Saiardi, A., Pisani, A., Baik, J.H., Centonze, D., Mercuri, N.B., Bernardi, G., and Borrelli, E. (1997). Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. J. Neurosci. *17*, 4536–4544.

Calabresi, P., Picconi, B., Tozzi, A., and Di Filippo, M. (2007). Dopamine-mediated regulation of corticostriatal synaptic plasticity. Trends Neurosci. *30*, 211–219.

Costa, R.M., Cohen, D., and Nicolelis, M.A.L. (2004). Differential corticostriatal plasticity during fast and slow motor skill learning in mice. Curr. Biol. *14*, 1124–1134.

Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., Lehéricy, S., and Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. Behav. Brain Res. *199*, 61–75.

El Yacoubi, M., Ledent, C., Ménard, J.F., Parmentier, M., Costentin, J., and Vaugeois, J.M. (2000). The stimulant effects of caffeine on locomotor behaviour in mice are mediated through its blockade of adenosine A(2A) receptors. Br. J. Pharmacol. *129*, 1465–1473.

Frank, M.J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. *17*, 51–72.

Frank, M.J., Seeberger, L.C., and O'reilly, R.C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science *306*, 1940–1943.

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

Gertler, T.S., Chan, C.S., and Surmeier, D.J. (2008). Dichotomous anatomical properties of adult striatal medium spiny neurons. J. Neurosci. *28*, 10814–10824.

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

Hikida, T., Kimura, K., Wada, N., Funabiki, K., and Nakanishi, S. (2010). Distinct roles of synaptic transmission in direct and indirect striatal pathways to reward and aversive behavior. Neuron *66*, 896–907.

Hornykiewicz, O. (2001). Chemical neuroanatomy of the basal ganglia—normal and in Parkinson's disease. J. Chem. Neuroanat. *22*, 3–12.

Hsu, C.W., Wang, C.S., and Chiu, T.H. (2010). Caffeine and a selective adenosine A2A receptor antagonist induce sensitization and cross-sensitization behavior associated with increased striatal dopamine in mice. J. Biomed. Sci. *17*, 4.

Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature *466*, 457–462.

Kelsey, J.E., Langelier, N.A., Oriel, B.S., and Reedy, C. (2009). The effects of systemic, intrastriatal, and intrapallidal injections of caffeine and systemic injections of A2A and A1 antagonists on forepaw stepping in the unilateral 6-OHDA-lesioned rat. Psychopharmacology (Berl.) *201*, 529–539.

Keus, S.H.J., Munneke, M., Nijkrake, M.J., Kwakkel, G., and Bloem, B.R. (2009). Physical therapy in Parkinson's disease: evolution and future challenges. Mov. Disord. *24*, 1–14.

Kostic, V.S., Svetel, M., Sternic, N., Dragasevic, N., and Przedborski, S. (1999). Theophylline increases "on" time in advanced parkinsonian patients. Neurology *52*, 1916.

Kravitz, A.V., Freeze, B.S., Parker, P.R.L., Kay, K., Thwin, M.T., Deisseroth, K., and Kreitzer, A.C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. Nature *466*, 622–626.

Kreitzer, A.C., and Malenka, R.C. (2007). Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. Nature *445*, 643–647.

Kreitzer, A.C., and Malenka, R.C. (2008). Striatal plasticity and basal ganglia circuit function. Neuron *60*, 543–554.

Kulisevsky, J., Barbanoj, M., Gironell, A., Antonijoan, R., Casas, M., and Pascual-Sedano, B. (2002). A double-blind crossover, placebo-controlled study of the adenosine A2A antagonist theophylline in Parkinson's disease. Clin. Neuropharmacol. *25*, 25–31.

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

Mallet, N., Le Moine, C., Charpier, S., and Gonon, F. (2005). Feedforward inhibition of projection neurons by fast-spiking GABA interneurons in the rat striatum in vivo. J. Neurosci. *25*, 3857–3869.

Mally, J., and Stone, T.W. (1996). Potential role of adenosine antagonist therapy in pathological tremor disorders. Pharmacol. Ther. *72*, 243–250.

Mink, J.W. (1996). The basal ganglia: focused selection and inhibition of competing motor programs. Prog. Neurobiol. *50*, 381–425.

Miyachi, S., Hikosaka, O., and Lu, X. (2002). Differential activation of monkey striatal neurons in the early and late stages of procedural learning. Exp. Brain Res. *146*, 122–126.

Muenter, M., and Tyce, G. (1971). L-dopa therapy of Parkinson's disease: plasma L-dopa concentration, therapeutic response, and side effects. Mayo Clin. Proc. *46*, 231–239.

Nieuwboer, A., Rochester, L., Müncks, L., and Swinnen, S.P. (2009). Motor learning in Parkinson's disease: limitations and potential for rehabilitation. Parkinsonism Relat. Disord. *15*(*Suppl 3*), S53–S58.

Nutt, J., Carter, J., and Van Houten, L. (1997). Short- and long-duration responses to levodopa during the first year of levodopa therapy. Ann. Neurol. *42*, 349–355.

Palminteri, S., Lebreton, M., Worbe, Y., Grabli, D., Hartmann, A., and Pessiglione, M. (2009). Pharmacological modulation of subliminal learning in Parkinson's and Tourette's syndromes. Proc. Natl. Acad. Sci. USA *106*, 19179–19184.

Peterson, J.D., Goldberg, J.A., and Surmeier, D.J. (2012). Adenosine A2a receptor antagonists attenuate striatal adaptations following dopamine depletion. Neurobiol. Dis. *45*, 409–416.

Picconi, B., Centonze, D., Håkansson, K., Bernardi, G., Greengard, P., Fisone, G., Cenci, M.A., and Calabresi, P. (2003). Loss of bidirectional striatal synaptic plasticity in L-DOPA-induced dyskinesia. Nat. Neurosci. *6*, 501–506.

Poldrack, R.A., Sabb, F.W., Foerde, K., Tom, S.M., Asarnow, R.F., Bookheimer, S.Y., and Knowlton, B.J. (2005). The neural correlates of motor skill automaticity. J. Neurosci. *25*, 5356–5364.

Puttemans, V., Wenderoth, N., and Swinnen, S.P. (2005). Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: from the cognitive stage to advanced levels of automaticity. J. Neurosci. *25*, 4270–4278.

Quiroz, C., Luján, R., Uchigashima, M., Simoes, A.P., Lerner, T.N., Borycz, J., Kachroo, A., Canas, P.M., Orru, M., Schwarzschild, M.A., et al. (2009). Key modulatory role of presynaptic adenosine A2A receptors in cortical neurotransmission to the striatal direct pathway. ScientificWorldJournal *9*, 1321–1344.

Schiffmann, S.N., Dassesse, D., d'Alcantara, P., Ledent, C., Swillens, S., and Zoli, M. (2003). A2A receptor and striatal cellular functions: regulation of gene expression, currents, and synaptic transmission. Neurology *61*(11, Suppl 6), S24–S29.

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D.J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. Science *321*, 848–851.

Surmeier, D.J., Plotkin, J., and Shen, W. (2009). Dopamine and synaptic plasticity in dorsal striatal circuits controlling action selection. Curr. Opin. Neurobiol. *19*, 621–628.

Tang, C., Pawlak, A.P., Prokopenko, V., and West, M.O. (2007). Changes in activity of the striatum during formation of a motor habit. Eur. J. Neurosci. *25*, 1212–1227.

Tozzi, A., Tscherter, A., Belcastro, V., Tantucci, M., Costa, C., Picconi, B., Centonze, D., Calabresi, P., and Borsini, F. (2007). Interaction of A2A adenosine and D2 dopamine receptors modulates corticostriatal glutamatergic transmission. Neuropharmacology *53*, 783–789.

Wickens, J.R. (2009). Synaptic plasticity in the basal ganglia. Behav. Brain Res. *199*, 119–128.

Wiecki, T.V., and Frank, M.J. (2010). Neurocomputational models of motor and cognitive deficits in Parkinson's disease. Prog. Brain Res. *183*, 275–297.

Wiecki, T.V., Riedinger, K., von Ameln-Mayerhofer, A., Schmidt, W.J., and Frank, M.J. (2009). A neurocomputational account of catalepsy sensitization induced by D2 receptor blockade in rats: context dependency, extinction, and renewal. Psychopharmacology (Berl.) *204*, 265–277.

Yin, H.H., Mulcare, S.P., Hilário, M.R.F., Clouse, E., Holloway, T., Davis, M.I., Hansson, A.C., Lovinger, D.M., and Costa, R.M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat. Neurosci. *12*, 333–341.

Cell PRESS
Open ACCESS

## EXTENDED EXPERIMENTAL PROCEDURES

### BG Model Implementation

The BG model can be obtained by emailing M.F. (michael_frank@brown.edu). Several demonstrations of reinforcement learning processes using this model are available for download at http://ski.clps.brown.edu/BG_Projects/.

The model is implemented with the emergent neural simulation software package (Aisa et al., 2008), adapted to simulate the anatomical and physiological properties of the BG circuitry in reinforcement learning and decision making (Frank, 2006). Emergent uses point neurons with excitatory, inhibitory, and leak conductances contributing to an integrated membrane potential, which is then thresholded and transformed to produce a rate code output communicated to other units. In the BG model, there is no supervised learning signal; reinforcement learning in the model relies on modification of corticostriatal synaptic strengths. Dopamine in the BG modifies activity in Go and NoGo units in the striatum, where this modulation of activity affects both the propensity for overall gating (Go relative to NoGo activity) and activity-dependent plasticity that occurs during reward prediction errors (Frank, 2005; Wiecki et al., 2009). Both of these functions are detailed below.

The below equations are written in general form; parameters vary according to physiological properties of different BG nuclei. For example, GPi/GPe units are tonically active in the absence of synaptic input, whereas striatal units fire only with convergent excitatory synaptic input from sensory input and preSMA. The below model neuron parameters are adjusted to capture these properties as described in Frank (2006).

The membrane potential $V_m$ is updated as a function of ionic conductances $g$ with reversal (driving) potentials $E$ according to the following differential equation:

$$\frac{dV_m(t)}{dt} = \tau \sum_c g_c(t)\bar{g}_c(E_c - V_m(t)), \tag{1}$$

with 4 channels ($c$) corresponding to: $e$ excitatory input; $l$ leak current; $i$ inhibitory input; $a$ accommodation. The reversal or equilibrium potentials $E_c$ determine the driving force of each of the channels, whereby $E_e$ is greater than the resting potential and $E_l$, $E_i$ and $E_a$ are typically less than resting potential (with the exception of tonically active neurons in GPi and GPe, where leak drives current into the neuron; Frank, 2006). Following electrophysiological convention, the overall conductance is decomposed into a time-varying component $g_c(t)$ computed as a function of the dynamic state of the network, and a constant $\bar{g}_c$ that controls the relative influence of the different conductances. The excitatory net input/conductance $g_e(t)$ or $\eta_j$ is computed as the proportion of open excitatory channels as a function of sending activations times the weight values:

$$\eta_j = g_e(t) = \langle x_i w_{ij} \rangle = \frac{1}{n} \sum_i x_i w_{ij}. \tag{2}$$

For units with inhibitory inputs from other layers (red projections in the model Figure in the main text), which are predominant in the basal ganglia, the inhibitory conductance is computed similarly, whereby $g_i(t)$ varies as a function of the sum of the synaptic inputs. Leak is a constant.

Dopamine adds an excitatory current to the Go units by modulating their excitability from cortical glutamatergic inputs, thereby simulating effects of D1 receptors. Dopamine also adds an inhibitory current to the NoGo units, simulating effects of D2 receptors.

Activation communicated to other cells ($y_j$) is a thresholded ($\Theta$) sigmoidal function of the membrane potential with gain parameter $\chi$:

$$y_j(t) = \frac{1}{\left(1 + \frac{1}{\gamma[V_m(t) - \Theta]_+}\right)}, \tag{3}$$

where $[x]_+$ is a threshold function that returns 0 if $x < 0$ and $x$ if $x > 0$. Note that if it returns 0, we assume $y_j(t) = 0$, to avoid dividing by 0. As it is, this function has a very sharp threshold, which interferes with graded learning mechanisms (e.g., gradient descent). To produce a less discontinuous deterministic function with a softer threshold, the function is convolved with a Gaussian noise kernel $(\mu = 0, \sigma = .005)$, which reflects the intrinsic processing noise of biological neurons:

$$y_j^*(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-x^2/(2\sigma^2)} y_j(z - x)dz, \tag{4}$$

where $x$ represents the $[V_m(t) - \Theta]_+$ value, and $y^*_j(x)$ is the noise-convolved activation for that value. In the simulation, this function is implemented using a numerical lookup table.

### Inhibition within and between Layers

Inhibition *between* layers (i.e., for GABAergic projections from Striatum to GPi/GPe, GPe to GPi and GPi to Thalamus) is achieved via unit inhibition, where the inhibitory current $g_i$ for the receiving unit is determined from the net input of the sending unit in the same way as described for $g_e$ (see above).

In the Striatum there are also fast-spiking GABAergic interneurons regulating the activity of both Go and NoGo units. These interneurons receive feedforward excitatory input from all areas projecting to Striatum. In addition, each Go unit receives a collateral inhibitory projection from units in the corresponding NoGo column. This allows the NoGo units to provide a "veto" function by suppressing Go unit activity when an action has been learned to be maladaptive, and is further motivated by electrophysiological data showing asymmetrical inhibition from NoGo to Go units (Taverna et al., 2008). Note that in the original BG model (Frank, 2005) inhibitory interneurons were not used and were simplified with a kWTA function (see below). However, this function cannot be used in instances where we wish to separately modulate D1 and D2 receptor strengths (in brief, because the D2 receptor projection is inhibitory and using the kWTA function does not allow for separate modulation of inhibitory strengths from effects other than local activity dynamics). Prior models investigating D2 strengths have also used inhibitory interneurons for this reason (in addition to the increased realism; Wiecki et al., 2009; Santesso et al., 2009).

There is also within layer lateral inhibition in the premotor cortex, leading to a competitive dynamic for motor response selection. However, for simplicity this inhibition is implemented with a kWTA (k-Winners-Take-All) function. The kWTA function computes a uniform level of inhibitory current for all units in the layer, such that the *k+1*th most excited unit within a layer is generally below its firing threshold, while the *k*th is typically above threshold. Thus, the kWTA function provides a computationally effective and efficient approximation to inhibitory dynamics which is typically used for cortical layers in leabra.

KWTA is computed via a uniform level of inhibitory current for all units in the layer as follows:

$$g_i = g^{\Theta}_{k+1} + q\left(g^{\Theta}_{k} - g^{\Theta}_{k+1}\right), \tag{5}$$

where $0 < q < 1$ (.25 default used here) is a parameter for setting the inhibition between the upper bound of $g^{\Theta}_{k}$ and the lower bound of $g^{\Theta}_{k+1}$. These boundary inhibition values are computed as a function of the level of inhibition necessary to keep a unit right at threshold:

$$g^{\Theta}_{i} = \frac{g^{*}_{e}\bar{g}_{e}(E_e - \Theta) + g_l\bar{g}_l(E_l - \Theta)}{\Theta - E_i}, \tag{6}$$

where $g^{*}_{e}$ is the excitatory net input.

## Connectivity

The connectivity of the BG network is critical, and is thus summarized here (see Frank, 2005, 2006 for details and references). Unless stated otherwise, projections depicted in the Figure are fully connected (that is all units from the source region target the destination region, with a randomly initialized synaptic weight matrix). However the units in PMC, Striatum, GPi, GPe, and Thalamus are all organized with columnar structure. Units in the first column of PMC represent one motor response and project to a single column of each of Go and NoGo units in the Striatum, which in turn project to the corresponding columns in GPi/GPe and Thalamus. GPe units project to their corresponding column in GPi. Each Thalamic unit is reciprocally connected with the associated column in PMC. This connectivity is similar to that described by anatomical studies, in which the same cortical region that projects to the striatum is modulated by the output through the BG circuitry and Thalamus (Kelly and Strick, 2004).

Dopamine units in the SNc project to the Striatum, with different projections to encode the effects of D1 receptors in Go neurons and D2 receptors in NoGo neurons. With increased dopamine, active Go units are excited while NoGo units are inhibited, and vice-versa with lowered dopamine levels. The particular set of units that are impacted by dopamine is determined by those receiving excitatory input from sensory cortex and PMC. Thus dopamine modulates this activity, thereby affecting the relative balance of Go versus NoGo activity in those units activated by cortex. This impact of dopamine on Go/NoGo activity levels influences both the propensity for gating (during response selection) and learning, as described next.

## Pharmacological Manipulations

In the main text we explored the effects of modulating the relative contributions of D1 and D2 receptors to performance and learning. The equation for computing excitatory conductances.

Equation 2 is subject to additional scaling of overall projections so that some 'count' more than others (meant to reflect the fact that neurons from some areas synapse closer or further to cell bodies and thereby have more or less influence on postsynaptic potentials than neurons from other areas, given the same synaptic weights). Specifically, excitatory conductances $g_{ek}$ arising as function of different projections k are scaled as follows:

$$g^{k}_{e} = s_k \frac{r_k}{\sum_p r_p} \left\langle x_i w_{ij} \right\rangle_k, \tag{7}$$

where $s_k$ provides an absolute scaling parameter for projection $k$ and $r_k$ reflects the relative scaling normalized by the sum of the scalings across all projections $p$. We simulated modulations of D1 and D2 receptor strengths by altering the scaling parameters $s_k$ for

these projections, where the defaults in the intact model were $s_{D1} = 0.3$ and $s_{D2} = 0.4$ (reflecting the increased affinity of D2 receptors for dopamine compared to D1 receptors). We then simulated the cocktail of D1+D2 blockade by reducing the impact of both of these projections, so that $s_{D1} = 0.2$ and $s_{D2} = 0.25$. To simulate D1 blockade alone, we found that greater changes were necessary to have an observable effect (much like in the mouse data, where higher doses were needed). We thus reduced the $s_{D1}$ to 0. For D2 blockade alone, we reduced $s_{D2}$ to 0.2.

## Learning

Synaptic connection weights in striatal units were learned using pure reinforcement learning (i.e., no supervised learning signal about which of the four responses the network should have made). In the *response phase*, the network settles into activity states based on input stimuli and its synaptic weights, ultimately gating one of the motor actions. In the *feedback phase*, the network resettles in the same manner, with the only difference being a change in simulated dopamine: an increase of SNc unit firing for positive reward prediction errors, and a decrease for negative prediction errors (Frank, 2005). This change in dopamine during the feedback phase modifies Go and NoGo activity levels which in turn affects plasticity, as described next.

For adjusting synaptic weights, the model uses a combination of Hebbian and contrastive Hebbian learning. The Hebbian term assumes simply that the level of activation of Go and NoGo units (and their presynaptic inputs) directly determines the synaptic weight change. The contrastive Hebbian component computes a simple difference of a pre and postsynaptic activation product across the response selection and outcome phases, which implies that learning occurs in proportion to the *change* in activation states from tonic to phasic dopamine levels. (Recall that dopamine influences Go versus NoGo activity levels by adding an excitatory current via simulated D1 dopamine receptors in Go units, and an inhibitory current via simulated D2 dopamine receptors in NoGo units. Thus increases in dopamine firing in SNc dopamine units promote active Go units to become more active, and NoGo units to become less active; vice-versa for pauses in dopamine).

The equation for the Hebbian weight change is:

$$\Delta_{hebb} w_{ij} = x_i^+ y_j^+ - y_j^+ w_{ij} = y_j^+ \left(x_i^+ - w_{ij}\right), \tag{8}$$

and for contrastive Hebbian learning:

$$\Delta_{err} w_{ij} = x_i^+ y_j^+ - x_i^- y_j^-, \tag{9}$$

which is subject to a soft-weight bounding to keep within the 0-1 range:

$$\Delta_{sbCHL} w_{ij} = [\Delta_{CHL}]_+ (1 - w_{ij}) + [\Delta_{CHL}]_- w_{ij}. \tag{10}$$

The two terms are then combined additively with a normalized mixing constant $k_{hebb}$:

$$\Delta w_{ij} = \varepsilon [k_{hebb}(\Delta_{hebb}) + (1 - k_{hebb})(\Delta_{sbCHL})]. \tag{11}$$

For projections to the Go units, $k_{hebb} = 0.01$ as is commonly used (default value in leabra) whereas projections to the NoGo units had $k_{hebb} = 0.1$, implying a stronger importance of Hebbian learning compared to contrastive Hebbian learning than usual. We confirmed that the qualitative patter of the data was robust to these assumptions and held when using similar levels of hebbian learning across inputs. However, note that the effects of D2 blockade are to enhance the excitability of NoGo units. Aberrant learning occurs when these units are active when they shouldn't be. Thus after the network selects a response, even if it is a correct response, NoGo units are active and exhibit potentiation. This effect is less prevalent with lower amounts of Hebbian learning. This same value of Hebbian learning was also used in the prior simulations of D2 blockade on catalepsy sensitization (Wiecki et al., 2009).

Learning rate $\varepsilon$ was set to 0.02, though again results were robust to other values. We simulated decrease in plasticity selectively to the NoGo units to capture A2A blockade by varying the learning rate to 0.01 and 0 (simulating low and higher doses).

## Reaction Times

As previously (Frank et al., 2007; Wiecki et al., 2009; Ratcliff and Frank, 2012), network reaction times are defined as the number of processing cycles until a motor response is gated by the thalamus (activation of a given thalamic unit reaches 50% maximal firing rate, but because this activity is ballistic once gating occurs the precise value is not critical).

## Simulating Acceleration in the Rotarod

To capture the accelerating property of the rotarod, we simply varied the maximum number of processing cycles (membrane potential updates) during which networks could select a response. Specifically, we initialized this value to 200 cycles. After each correct response (i.e., the model did not 'fall off' the rod). we decreased this value by 15 cycles. Thus with increasing correct actions, the demands on fast response execution were greater. After a single error, we reinitialized the time limit to 200. Once again, we confirmed that the general qualitative pattern of results were not dependent on this implementation and held without simulating the acceleration component, however this increased demand makes the networks more susceptible to errors and performance effects (e.g., D1 blockade after learning).

## SUPPLEMENTAL REFERENCES

Aisa, B., Mingus, B., and O'Reilly, R. (2008). The emergent neural modeling system. Neural. Netw. *21*, 1146–1152.

Frank, M.J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J. Cogn. Neurosci. *17*, 51–72.

Frank, M.J. (2006). Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw. *19*, 1120–1136.

Frank, M.J., Scheres, A., and Sherman, S.J. (2007). Understanding decision making deficits in neurological conditions: insights from models of natural action selection. Phil. Trans. R. Soc. B. *362*, 1641–1654.

Kelly, R.M., and Strick, P.L. (2004). Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits. Prog. Brain Res. *143*, 449–459.

Ratcliff, R., and Frank, M.J. (2012). Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models. Neural Comput. *24*, 1186–1229.

Santesso, D., Evins, A., Frank, M., Cowman, E., and Pizzagalli, D. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: evidence from event-related potentials and computational modeling of striatal-cortical function. Hum. Brain Mapp. *30*, 1963–1976.

Taverna, S., Ilijic, E., and Surmeier, D.J. (2008). Recurrent collateral connections of striatal medium spiny neurons are disrupted in models of Parkinson's disease. J. Neurosci. *28*, 5504–5512.

Wiecki, T.V., Riedinger, K., von Ameln-Mayerhofer, A., Schmidt, W.J., and Frank, M.J. (2009). A neurocomputational account of catalepsy sensitization induced by D2 receptor blockade in rats: context dependency, extinction, and renewal. Psychopharmacology (Berl.) *204*, 265–277.