Reward and punishment illuminated

Joseph J Paton & Kenway Louie

How do outcomes affect future behavior? A study using precise optogenetic stimulation finds that learning from positive reinforcement is mediated by striatal pathways distinct from those that mediate learning from punishment.

When a new cafe opens, how do we decide to stop for a cup of coffee? Absent any pre-existing knowledge, we have to taste a coffee or two and use that information to guide our morning routine. One hallmark of adaptive behavior is the ability to learn from the outcomes of actions, whether those results are positive or negative. This importance of outcomes in shaping behavior is codified in Thorndike’s law of effect: behaviors associated with satisfaction or discomfort will be more or less likely to recur, respectively. Decades of lesion, electrophysiology and functional imaging studies have outlined a broad network of brain areas involved in learning about reward and punishment, but how this processing is integrated with action selection remains unknown. In this issue of *Neuroscience*, Kravitz et al. report that two distinct pathways in the basal ganglia, a subcortical system that is essential for motor control, differentially mediate reward and punishment.

A critical tool for examining the neural basis of reinforcement has been the targeted activation of specific brain areas. Electrical stimulation has long been known to elicit behavioral phenomena, ranging from simple percepts and movements to emotions and vivid memories, depending on the targeted brain region. In learning protocols, stimulation itself can serve as reinforcement, driving animals to seek or avoid further stimulation. A wealth of brain stimulation reward studies have identified a network of subcortical areas involved in reward processing, notably the lateral hypothalamus, medial forebrain bundle and mesolimbic dopaminergic system. The firing of dopaminergic neurons may be critical for learning, signaling a reward prediction error representing whether the moment at hand is better or worse than expected. An intuition behind this kind of teaching signal is that if predictions turn out to be inaccurate in a given situation, this is a useful sign that the system should update its valuations to improve future predictions.

One substantial drawback of electrical stimulation, however, is non-specific activation in the vicinity of the stimulating electrode, including all local neuronal cell bodies and neuronal axons passing by en route to distant locations. Such non-specificity is particularly problematic for studying areas without a gross functional architecture; for example, nuclei where neurons projecting to different targets or encoding different information are closely intermingled. The recent arrival of optogenetic techniques, which combine light-activated spiking activity with genetic localization to specific cell types, represents a substantial advance in the specificity of targeted activation. Kravitz et al. used optogenetic control to selectively target two distinct populations of neurons in the striatum, the major input nucleus of the basal ganglia.

Roughly 95% of the neurons in the striatum are GABAergic projection cells called medium spiny neurons (MSNs), so named for their medium size and the dense distribution of spines along their dendrites. MSNs receive excitatory input from layer 5 pyramidal neurons in almost all areas of cortex, as well as from the thalamus, and modulatory input from dopaminergic neurons in the substantia nigrapars compacta and/or the ventral tegmental area in the midbrain. Striatal MSNs can be subdivided into two subclasses on the basis of projection target and expression of dopamine receptor types. Direct pathway striatonigral MSNs (dMSNs) express D1 dopamine receptors and project directly to basal ganglia output nuclei: the internal segment of the globus pallidus (GPI) and/or the substantia nigra pars reticulata (SNr). Indirect pathway striatopallidal MSNs (iMSNs) express D2 dopamine receptors and represent the first stage of a more indirect route to basal ganglia output, terminating primarily in the external segment of the globus pallidus (GPe).

The direct and indirect pathways comprise the two fundamental opposing forces in the classic model of basal ganglia motor control. Excitatory corticostriatal input to direct pathway dMSNs increases striatonigral inhibition of GPe and SNr activity, disinhibiting thalamocortical projections and facilitating movement (Fig. 1a). In the indirect pathway, striatal activity in iMSNs acting through the GPe and subthalamic nucleus results in a net inhibition of thalamocortical activity and suppression of movement (Fig. 1b). Action selection is thought to be implemented by a striatal competition between actions specified by corticostriatal inputs, mediated by balanced direct and indirect pathway activity. Although recent anatomical and functional evidence suggests that this scheme may be oversimplified, major motor pathologies can be explained by an imbalance in striatal information processing. In Parkinson’s disease, loss of dopaminergic input leads to an overactivity of indirect versus direct pathway activity, resulting in a poverty of movement; in Huntington’s disease, loss of indirect pathway activity removes inhibitory control, resulting in an excess of abnormal movements.

In addition to action selection, growing evidence also implicates the basal ganglia in reward learning. Theoretical models suggest that learning from positive and negative outcomes could be functionally segregated via the direct and indirect pathways, suggesting that reward and punishment may be mediated by separate anatomic systems. Given its reward-related signaling and dense innervation of the striatum, dopaminergic input is likely to be critical. Indeed, D1 type and D2 type receptors in the striatum have opposite effects on cell excitability: activation of D1 receptors increases excitability of dMSNs and activation of D2 receptors decreases excitability of iMSNs, both acting via G protein–coupled changes in responsiveness to glutamatergic input. The phasic release of dopamine around better-than-expected behavioral events would be expected...
Mice seek self-stimulation of direct pathway: reward

Mice avoid self-stimulation of indirect pathway: punishment

**Figure 1** Simplified schematic of basal ganglia anatomy and circuit response to optogenetic stimulation. (a,b) Kravitz *et al.* expressed the light-sensitive cation channel Channelrhodopsin2 (ChR2) in dMSNs (a) and iMSNs (b) by injecting adeno-associated virus (AAV) containing a loxP-flanked inverted ChR2 construct into the dorsomedial striatum of transgenic mice expressing Cre recombinase in dMSNs (D1-Cre) or iMSNs (A2A-Cre), respectively. They then placed the mice in an environment with two contact devices. One contact triggered a brief pulse of laser light delivered to the striatum through an optical fiber. The other contact was inactive. Activation of MSNs is thought to change the activity at many points in the downstream circuitry, ultimately leading to a decrease or increase in inhibitory output from the basal ganglia when stimulating dMSNs (a) or iMSNs (b), respectively. Arrow thickness indicates predicted relative activity in different projection pathways in response to the two stimulation conditions. Blue, direct pathway; red, indirect pathway. STN, subthalamic nucleus. Areas outlined in black send inhibitory projections. Those outlined in white send excitatory projections.

A critical question is how these reinforcement effects are related to the well-documented role of the basal ganglia in motor control. Previously, in an experiment using the same optogenetic techniques and mouse lines, direct pathway stimulation increased locomotion and decreased freezing, whereas indirect pathway stimulation produced a parkinsonian state, with decreased locomotion, increased freezing and bradykinesia. Such broad stimulation-induced movement changes raise the possibility that the behaviors observed in the current study arise from simple changes in overall motor activity. However, Kravitz *et al.* found that striatal activation induced few changes in movement parameters, possibly owing to short laser stimulation durations. Furthermore, although trigger biases emerged gradually over the course of initial training sessions, they were immediately observed in subsequent sessions, consistent with a learned behavior rather than a motor confound.

Which neural systems mediate this simulation-induced reinforcement learning? In the standard model of basal ganglia function, the rich dopaminergic innervation of the striatum can influence action selection in two primary ways. First, dopamine indirectly excites dMSNs and inhibits iMSNs, primarily by altering the responsiveness to glutamatergic input. Second, phasic dopamine bursts may induce long-lasting changes in striatal plasticity, including long-term potentiation via D1 receptor activation and long-term depression via D2 receptor activation. Notably, by targeting dopamine receptor–expressing striatal MSNs, Kravitz *et al.* effectively bypassed dopaminergic signaling. To confirm that their reinforcement effects were independent of dopamine, the authors repeated the operant conditioning experiments after injecting a combination of D1 and D2 receptor antagonists. Although these antagonists reduced overall locomotor activity, the positive and negative trigger biases associated with direct and indirect pathway stimulation remained intact.

These findings suggest that the activation of striatal pathways can mediate the effects of reinforcement independent of dopaminergic signaling, but the site of learning remains an open question. Although learning appears to be independent of dopamine-induced plasticity at synapses on MSNs, co-occurrence of striatal input and stimulation might change efficacy at corticostriatal and thalamostriatal synapses through other, dopamine-independent mechanisms, such as glutamate...
Overall, the results of Kravitz et al. highlight a fundamental point about decision-making: selecting an action is never truly independent of reward learning. Functionally, the learned values of different options is a crucial element of the action selection process. Neurally, action selection and reinforcement learning appear to be implemented in the same striatal circuitry, with distinct functional compartments processing rewarding versus aversive outcomes. Understanding the exact nature and mechanism of this relationship between reinforcement and action will be a critical avenue for further research.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Distinct roles for direct and indirect pathway striatal neurons in reinforcement

Alexxai V Kravitz, Lynne D Tye & Anatol C Kreitzer

Dopamine signaling is implicated in reinforcement learning, but the neural substrates targeted by dopamine are poorly understood. We bypassed dopamine signaling itself and tested how optogenetic activation of dopamine D1 or D2 receptor–expressing striatal projection neurons influenced reinforcement learning in mice. Stimulating D1 receptor–expressing neurons induced persistent reinforcement, whereas stimulating D2 receptor–expressing neurons induced transient punishment, indicating that activation of these circuits is sufficient to modify the probability of performing future actions.

Reinforcement and punishment are fundamental processes that shape animal learning. Reinforcement maintains or increases, whereas punishment decreases, the probability of specific behavior. Dysfunction in these processes contributes to many psychiatric disorders. For example, addiction is characterized by heightened reinforcement from drug-paired stimuli, coupled with impaired punishment from negative consequences. In contrast, depression is marked by impaired reinforcement from positive stimuli and heightened punishment from negative stimuli. Although the striatum has been implicated in both reinforcement and punishment, the specific roles of the two populations of striatal projection neurons are not well understood. We tested the hypothesis that D1 receptor–expressing direct pathway medium spiny neurons (dMSNs) mediate reinforcement and D2 receptor–expressing indirect pathway neurons (iMSNs) mediate punishment.

To selectively activate dMSNs or iMSNs in vivo, we expressed channelrhodopsin-2 (ChR2) in the dorsomedial striatum using a Cre-dependent viral strategy. We first characterized the effects of ChR2 stimulation on dMSNs in awake behaving mice with in vivo electrophysiology, using microwire arrays that included an integrated optical fiber (Supplementary Fig. 1a,b). Each of 48 recorded neurons (n = 3 mice) was illuminated at four laser intensities (0.1, 0.3, 1 and 3 mW, 1-s constant illumination). We concluded that neurons expressed ChR2 (and were therefore dMSNs) if they exhibited a significant increase in firing in 40 ms of the laser onset at any laser power (Supplementary Fig. 1c–g). Spiking data from recorded neurons in this experiment are available for download at http://uri.neuinfo.org/nif/nifstd/nif_144028. Overall, 19 (40%) neurons were identified as dMSNs. Notably, there were no significant differences between waveform characteristics or sorting quality of recording channels that contained ChR2-positive MSNs versus those that did not (P > 0.30 for all tests; Supplementary Table 1). Average firing rates and total number of ChR2-responsive dMSNs increased with higher laser intensity, indicating that higher laser intensity caused more MSN activation (Supplementary Fig. 1h–j).

Figure 1 dMSN stimulation induces persistent reinforcement, whereas iMSN stimulation induces transient punishment. (a,b) Percent of contacts with the laser-paired trigger across each session for dMSN-ChR2 (blue) and iMSN-ChR2 (red) mice (2-min bins). (c,d) Percent of contacts with the laser-paired trigger during retraining and extinction sessions (2-min bins). (e,f) Probability of subsequent contact with the laser-paired trigger following a previous trigger activation for dMSN-ChR2, YFP control and iMSN-ChR2 mice. (g) Peri-event time histogram of velocity for dMSN-ChR2 and iMSN-ChR2 mice following contact with the laser-paired (blue) or inactive (gray) triggers. The time of stimulation for the laser-paired trigger is shown in blue. Error bars in all panels represent s.e.m.

* represents significant differences and NS represents nonsignificant differences, based on an alpha of 0.05.

1 Gladstone Institute of Neurological Disease, University of California San Francisco, San Francisco, California, USA. 2 Departments of Physiology and Neurology, University of California San Francisco, San Francisco, California, USA. 3 Neuroscience Graduate Program, University of California San Francisco, San Francisco, California, USA. 4 These authors contributed equally to this work. Correspondence should be addressed to A.C.K. (akreitzer@gladstone.ucsf.edu).
To investigate reinforcement, we gave mice expressing ChR2 in dMSNs or iMSNs (termed dMSN-ChR2 or iMSN-ChR2 mice, respectively) bilateral fiber optic implants targeting the dorsomedial striatum (Supplementary Fig. 2). The dorsomedial striatum was targeted because of its role in reinforcement and action selection.\(^5,8,9,11\) Mice were placed in an operant box that contained two capacitive touch triggers, one that activated a 1-mW laser (1-s constant illumination, delivered bilaterally) and one that was inactive (Supplementary Video 1). The triggers were crucial for this experiment because they are much more sensitive than lever-press or nose-poke manipulations, which allowed us to observe both increments and decrements in responding. We tested three groups of naive mice in this task: dMSN-ChR2 mice (n = 8), iMSN-ChR2 mice (n = 8) and control mice that expressed yellow fluorescent protein (YFP) in dMSNs (n = 4) or iMSNs (n = 4). Data from dMSN-YFP and iMSN-YFP control mice were combined, as neither group showed any significant effects (all P > 0.25 when the below analyses were run on each group independently; Supplementary Fig. 3 and Supplementary Table 2). All groups completed one 30-min training session each day for three consecutive days.

In the first session, naive dMSN-ChR2 mice exhibited a significant bias toward the laser-paired trigger (P < 0.001), whereas iMSN-ChR2 mice exhibited a significant bias away from the laser-paired trigger (P < 0.0001; Fig. 1a, Supplementary Table 2 and Supplementary Videos 1 and 2). In contrast with the first 2 min of day 1, trigger biases for both dMSN-ChR2 and iMSN-ChR2 mice were present in the first 2 min of day 2, suggesting a learned behavior (Fig. 1b). However, this effect was weaker in iMSN-ChR2 mice and was no longer significant at the beginning of day 3 (P = 0.33; Fig. 1b). To further investigate this persistence, we subjected mice with at least 3 d of prior training to 30 min of retraining, followed by a 30-min extinction session, in which neither trigger elicited a laser pulse. dMSN-ChR2 mice continued to exhibit a significant bias toward the previously laser-paired trigger throughout the entire extinction session (P < 0.01; Supplementary Table 2), whereas iMSN-ChR2 mice rapidly lost their behavioral preference (Fig. 1c,d and Supplementary Table 2).

In light of these differences in persistence, we analyzed the time course of reinforcement and punishment following each laser pulse and noted differences on a shorter time scale. dMSN-ChR2 mice had a heightened probability of contacting the laser-paired trigger for at least 45 s following a laser pulse relative to YFP control mice (P < 0.001; Fig. 1e,f). iMSN-ChR2 mice had a lower probability of contacting the laser-paired trigger in the initial 15 s following a laser pulse (P < 0.01), but this effect was no longer significant in the 15–30 s following the laser stimulation (P = 0.40; Fig. 1e,f). These findings are consistent with the diminished cross-day persistence of iMSN-mediated punishment. We considered the possibility that the amount of experience dMSN-ChR2 and iMSN-ChR2 mice had with the laser could explain these differences in learning. However, the persistence of trigger preference across sessions was not related to the number of contacts mice had with the laser-paired trigger on the previous day (Supplementary Fig. 4).

As activation of these cell groups can induce motor changes,\(^10\) we tested whether motoric changes during the laser pulses might have contributed to our results. For example, dMSN-ChR2 activation might have induced stereotypes that caused multiple contacts during the laser-paired stimulation. Notably, however, dMSN-ChR2 stimulation did not produce changes in the animal’s velocity (Fig. 1g and Supplementary Video 1), which is different from what we saw in a previous study,\(^10\) which may reflect either the shorter duration...
(1 versus 30 s) or the operant nature of this stimulation. iMSN-ChR2 stimulation elicited brief freezing (consistent with our previous findings) followed by an aversive-like escape response, evidenced by an increase in velocity following the laser pulse (Fig. 1g and Supplementary Video 2). However, these brief (<2 s) changes in motor behavior following stimulation are not sufficient to explain the decrease in probability of active trigger contacts that persisted for >15 s after stimulation (Fig. 1e).

To test whether the level of DMSN activation correlated with the magnitude of reinforcement, we placed dMSN-ChR2 mice (same cohort as shown in Fig. 1; also see Supplementary Fig. 2a) in an operant box that used four capacitive touch sensors as operant triggers. A computer detected contacts with these triggers and controlled three lasers, which were calibrated to 0.3, 1 and 3 mW of output power per side (1 s constant illumination, delivered bilaterally; Supplementary Fig. 5a). Contacts with an inactive trigger were also counted, but had no consequences. dMSN-ChR2 mice preferred higher laser intensities (R2 = 0.98, P < 0.01, n = 8; Supplementary Fig. 5b,c), demonstrating that the magnitude of reinforcement was correlated with the level of dMSN activation.

Although we were directly activating MSNs, we considered the possibility that we might have also elicited striatal dopamine (DA) release. To examine whether DA itself was involved in the acquisition of trigger preference, we tested whether combined D1 and D2 receptor antagonists (0.02 mg per kg of body weight SCH23390 and 25 mg per kg sulphiride, co-injected intraperitoneally) would impair the acquisition of the two-trigger operant task in naive dMSN-ChR2 and iMSN-ChR2 mice (Fig. 2a and Supplementary Fig. 2b). DA antagonists significantly reduced overall movement compared with that in separate groups of mice that were injected with saline (P < 0.01; Fig. 2a). Notably, DA antagonists did not significantly alter the total number of contacts with either trigger (P = 0.25; Fig. 2a) or prevent acquisition of trigger biases over 3 d of training (Fig. 2b,c and Supplementary Table 3). To test whether DA was required for the expression of trigger bias, we injected the previously saline-treated groups with the same DA antagonists on a fourth day of training and found that expression of the previously learned trigger preference was not impaired (Fig. 2d and Supplementary Table 3).

To test whether this learning was specific to our operant task, we trained dMSN-ChR2 and iMSN-ChR2 mice in a real-time place-preference task in which one-half of a chamber was paired with pulsed laser stimulation (2 s of 1-mW laser and 8 s off; cohort is a subset of mice shown in Fig. 2). Mice were trained for 30 min for two consecutive days, and the second training session was immediately followed by a 30-min test session with no laser stimulation (Fig. 3a). Consistent with our results in the operant task, dMSN-ChR2 mice showed a persistence of their learned place preference during the entire test session, whereas iMSN-ChR2 mice showed no evidence of such persistence (Fig. 3).

Our results indicate that activation of striatal dMSNs is sufficient for persistent reinforcement, whereas activation of iMSNs is sufficient for transient punishment, in both an operant and a place-preference task. The differences in time course that we observed are qualitatively similar to results from animals as diverse as invertebrates, rodents and humans, indicating that reinforcement is more effective than punishment at modifying long-term behavior. These differences in time course may relate to differences in synaptic plasticity mechanisms in each pathway. Although DA is known to influence both activity and plasticity of these cells under natural conditions, other neurochemicals are also important. Future therapies could target dMSNs or iMSNs independently to address specific dysfunctions in reinforcement or punishment associated with psychiatric disorders.

**METHODS**

Methods and any associated references are available in the online version of the paper.

*Note: Supplementary information is available in the online version of the paper.*

**ACKNOWLEDGMENTS**

We thank the Nikon Imaging Center at the University of California San Francisco for assistance with image acquisition. K. Deisseroth for optogenetic constructs and K. Tye for helpful comments on the manuscript. A.C.K. and co-workers are funded by the W.M. Keck Foundation, the Pew Biomedical Scholars Program, the McKnight Foundation and the US National Institutes of Health.

**AUTHOR CONTRIBUTIONS**

A.V.K. and L.D.T. jointly conducted the experiments and analyzed the data. A.V.K. and A.C.K. conceived the study and wrote the manuscript.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/nn.3100. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

Viral expression of DIO-ChR2-YFP and DIO-YFP. We used double loxP-flanked inverted (DIO) constructs to express ChR2-YFP fusions and YFP alone inCre-expressing neurons, which virtually eliminates recombination in cells that do not express Cre recombinase.16 The double loxP-flanked reverse ChR2-YFP or YFP cassette was cloned into a modified version of the pAAV2-MCS vector (Stratagene) carrying the EEF1A1 promoter and the Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) to enhance expression. The recombinant AAV vectors were serotyped with AAV1 coat proteins and packaged by the viral vector core at the University of North Carolina. The final viral concentration was 4 × 10^12 virus molecules ml^−1 (by Dot Blot, University of North Carolina vector core).

Stereotaxic viral injections and implantation of fiber optic ferrules. Anesthesia was induced with a mixture of ketamine and xylazine (100 mg ketamine and 5 mg xylazine per kg body weight, co-injected intraperitoneally), and maintained with 0.5–1.0% isoflurane (vol/vol) through a nose cone mounted on a stereotaxic apparatus (Kopf Instruments). The scalp was opened and bilateral holes were drilled in the skull (+0.8 mm anterior and ±1.5 mm lateral from bregma). We injected 1 µl of DIO ChR2-YFP virus into the left and right dorsomedial striata –3.0 mm below the top of the skull from top of brain through a 33 gauge steel injector cannula (Plastics1) using a syringe pump (World Precision Instruments) over 10 min. The injection cannula was left in place for 5 min following the injection and then slowly removed. After the viral injection, a plastic mount containing two fibers (105-µm core and 125-µm cladding) mounted in 1.25-mm zirconia ferrules were slowly lowered into the brain and cemented in place such that each fiber was aimed at the dorsomedial striatum on either side. To allow time for viral expression, we housed animals for at least 2 weeks following injection before any experiments were initiated. All surgical procedures were performed under aseptic conditions.

Implantation of electrode arrays for awake recordings. Anesthesia was induced with a mixture of ketamine and xylazine (100 mg ketamine and 5 mg xylazine per kg body weight, co-injected intraperitoneally) and maintained with isoflurane through a nose cone mounted on a stereotaxic apparatus (Kopf Instruments). The scalp was opened and a hole was drilled in the skull (0.0 to +1.0 mm AP, −1.0 to −2.0 mm ML from bregma). DIO ChR2-YFP virus was injected into this hole as described above. Two skull screws were implanted in the opposing hemisphere. Dental adhesive (C&B Metabond, Parkell) was used to fix the skull screws in place and coat the surface of the skull. An array of 16 or 32 microwires (35-µm tungsten wires, 100-µm spacing between wires, 200-µm spacing between rows, Innovative Physiology) and one optical fiber in a ferrule was lowered into the striatum (3.0 mm below the surface of the skull) and cemented in place with dental acrylic (Ortho-Jet, Lang Dental). After the cement dried, the array was sutured shut. Animals were allowed to recover for at least 2 weeks before striatal recordings were made.

**In vivo electrophysiology.** Voltage signals from each recording site on the microwire array were bandpass-filtered, such that activity between 150 and 8,000 Hz was analyzed as spiking activity. This data was amplified, processed and digitally captured using commercial hardware and software (Plexon). Single units were discriminated with principal component analysis (Offline sorter, Plexon). Two criteria were used to ensure quality of recorded units: recorded units smaller than 100 µV (−3 times the noise band) were excluded from further analysis and recorded units in which more than 1% of interspike intervals were shorter than 2 ms were excluded from further analysis. Average waveforms were exported with Offline sorter. During the recording we coupled the array to a laser and pulsed the laser at four intensities (0.1, 0.3, 1 and 3 mW). Laser stimulation was run in a cyclical fashion, on for 1 s and off for 3 s. Each neuron received 100 pulses at each laser intensity.

**Identification of ChR2-expressing units in in vivo recordings.** For all neurons, peri-event histograms were generated for each laser intensity independently. Neurons were classified as ChR2 expressing if they exhibited, within 40 ms of the laser onset, a firing rate more than threefold greater than the s.d. of the s.d. of the s.d. preceding the laser pulse. Each neuron was tested independently at each laser power, and neurons that satisfied these criteria at any one power were defined as ChR2-expressing cells.

**Behavioral experiments.** Two 1-m glass fibers (62.5-µm core, 125-µm cladding, Ecablarmat.com) were connectorized with LC connectors on one end, and an LC ferrule on the other. The LC connectorized ends were hooked up to a 50/50 splitter coming off a laser, such that equal laser light passed through each fiber. The end with the LC ferrule was attached to the mouse for experiments with a zirconia connection sleeve. After the optical fibers were connected, each mouse was placed in a 16-inch square operant box that contained either two or four capacitive touch sensors, which were used as operant triggers. Contacts with the touch sensors were recorded by Ethovision 7.0 software, which controlled the illumination of lasers (1 s constant illumination, delivered bilaterally) via an I/O box (Noldus). A separate behavioral chamber was used for the place preference task, which contained two 8-× 8-inch compartments, one of which had opaque white walls and the other opaque black walls. The mouse’s position was calculated in real time with the Ethovision 7.0 software, and this position was used to control the illumination of the laser, which was counterbalanced between the white and black compartments in each group. Laser was illuminated in a 2-s on and 8-s off cycle for the duration that the mouse remained in the laser-paired compartment. For drug experiments, mice were injected with DA antagonists (0.02 mg per kg SCH23390 and sulpiride 25 mg per kg co-injected intraperitoneally) or 0.9% saline (wt/vol) and returned to their home cage for 30 min before beginning each 30-min session. All experimental sessions were 30 min long.

**Statistical analyses.** Statistics were first performed with repeated-measures ANOVAs to test for effects of day (day 1, 2 or 3), group (dMSN, iMSN or YFP) and/or drug (antagonist or saline), followed up with post hoc one-sample t tests to test whether specific conditions differed from the null hypothesis that 50% of behavior would be directed at the laser-paired and 50% percent at the inactive trigger or compartment.

**Histology.** Animals were killed with a lethal dose of ketamine and xylazine (400 mg ketamine + 20 mg xylazine per kg body weight, intraperitoneal). Animals with microwire arrays received a current injection (10 µA for 5 s) through each microwire to lesion the wire tips. All animals were transcardially perfused with phosphate-buffered saline, followed by 4% paraformaldehyde (wt/vol). Following perfusion, brains were left in 4% paraformaldehyde for 16–24 h and then moved to a 30% sucrose solution (wt/vol) in phosphate-buffered saline for 2–3 d. Brains were then frozen and cut into 30-µm sections (either coronal or sagittal) with a sliding microtome (Leica Microsystems, model SM2000R) equipped with a freezing stage (Physiotemp). To identify fiber locations, relevant sections were identified and mounted on slides. Sections were then photographed in bright field and fluorescence on a Nikon D6 microscope with a 4× objective. From these photographs, fiber tip locations were identified and marked on a coronal schematic of the striatum at 0.8 mm anterior to bregma.