nature neuroscience

Distinct roles for direct and indirect pathway striatal neurons in reinforcement

Alexxai V Kravitz[1,4], Lynne D Tye[1,2,4] & Anatol C Kreitzer[1–3]

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[1,2]. 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[3]. In contrast, depression is marked by impaired reinforcement from positive stimuli and heightened punishment from negative stimuli[4]. 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[5–9].

To selectively activate dMSNs or iMSNs *in vivo*, we expressed channelrhodopsin-2 (ChR2) in the dorsomedial striatum using a Cre-dependent viral strategy[10]. 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/nlx_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**).

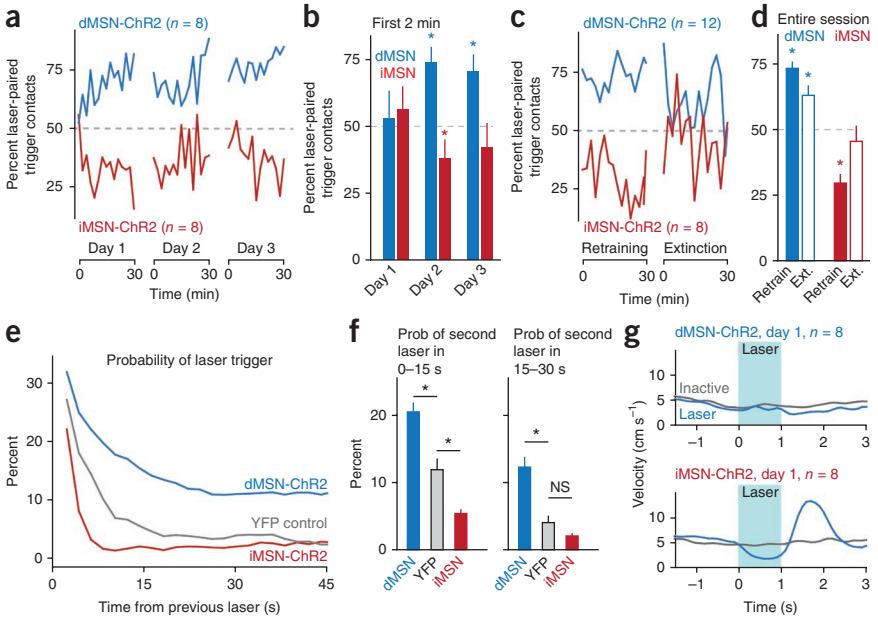

Figure 1 dMSN stimulation induces persistent reinforcement, whereas iMSN stimulation induces transient punishment. (**a**,**b**) Percent of contacts with the laser-paired trigger across each session for dMSN-ChR2 (blue) and iMSN-ChR2 (red) mice (2-min bins). (**c**,**d**) Percent of contacts with the 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).

Received 9 February; accepted 2 April; published online 29 April 2012; doi:10.1038/nn.3100

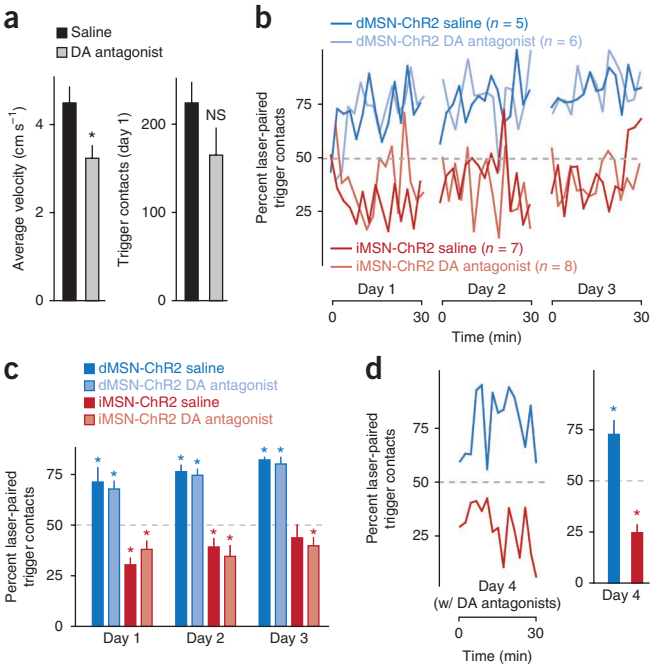

Figure 2 Acquisition and expression of trigger preference are not influenced by dopaminergic antagonists. (**a**) Average velocity and total number of trigger contacts for mice treated with saline (black bars) and DA antagonists (gray bars). (**b**) Acquisition of trigger preference for saline- (dark lines) and DA antagonist–treated mice (light lines) (days 1–3, 2-min bins) for dMSN-ChR2 (blue) and iMSN-ChR2 (red) mice. (**c**) Quantification of trigger preference for the entire duration of each session. (**d**) Expression of trigger preference in the saline-treated mice from **c** in a subsequent session in which mice were administered DA antagonists (day 4, 2-min bins). Error bars in all panels represent s.e.m. * represents significant differences and NS represents nonsignificant differences, based on an alpha of 0.05.

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 manipulanda, 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

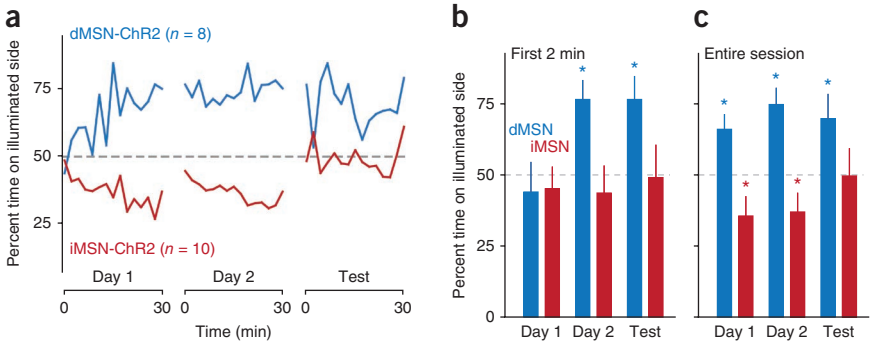

Figure 3 dMSN and iMSN stimulation modify place preference. (**a**) Percent of time animal spent in laser-paired compartment for dMSN-ChR2 (blue) and iMSN-ChR2 (red) mice (2-min bins) for two training days (laser illumination when the animal is in the laser-paired chamber) and for a test day (no laser illumination in either chamber). (**b**,**c**) Quantification of place preference during the first two min of each session (**b**) and for the entire duration of each session (**c**) for dMSN-ChR2 and iMSN-ChR2 mice. Error bars in all panels represent s.e.m. * represents significant differences, based on an alpha of 0.05.

(1 versus 30 s) or the operant nature of this stimulation. iMSN-ChR2 stimulation elicited brief freezing (consistent with our previous findings[10]) 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 ($R^2 = 0.98$, $P < 0.01$, $n = 8$; **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 sulpiride, 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[2,12,13]. These differences in time course may relate to differences in synaptic plasticity mechanisms in each pathway[14]. Although DA is known to influence both activity and plasticity of these cells under natural conditions[5,9,15], other neurochemicals are important as well. 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.

1. Azrin, N.H. & Holz, W.C. Punishment. in *Operant Behavior: Areas of Research and Application* (ed. Honig, W.K.) 380–447 (Appleton-Century-Crofts, New York, 1966).
2. Skinner, B.F. *Science and Human Behavior* (Macmillan, New York, 1953).
3. Koob, G.F. & Volkow, N.D. *Neuropsychopharmacology* **35**, 217–238 (2010).
4. Eshel, N. & Roiser, J.P. *Biol. Psychiatry* **68**, 118–124 (2010).
5. Bromberg-Martin, E.S., Matsumoto, M. & Hikosaka, O. *Neuron* **68**, 815–834 (2010).
6. Lobo, M.K. *et al. Science* **330**, 385–390 (2010).
7. Ferguson, S.M. *et al. Nat. Neurosci.* **14**, 22–24 (2011).
8. Hikida, T., Kimura, K., Wada, N., Funabiki, K. & Nakanishi, S. *Neuron* **66**, 896–907 (2010).
9. Frank, M.J., Seeberger, L.C. & O'Reilly, R.C. *Science* **306**, 1940–1943 (2004).
10. Kravitz, A.V. *et al. Nature* **466**, 622–626 (2010).
11. Balleine, B.W., Delgado, M.R. & Hikosaka, O. *J. Neurosci.* **27**, 8161–8165 (2007).
12. Nakatani, Y. *et al. Neurobiol. Learn. Mem.* **92**, 370–380 (2009).
13. Abe, M. *et al. Curr. Biol.* **21**, 557–562 (2011).
14. Kreitzer, A.C. & Malenka, R.C. *Neuron* **60**, 543–554 (2008).
15. Schultz, W. *Annu. Rev. Neurosci.* **30**, 259–288 (2007).

ONLINE METHODS

Subjects. Bacterial artificial chromosome (BAC) transgenic mouse lines that express Cre recombinase under control of the dopamine D1 receptor and A2A receptor regulatory elements were obtained from GENSAT. Animals entered the study at ~6 weeks of age, weighing ~20 g. All procedures were approved by the University of California San Francisco Institutional Animal Care and Use Committee.

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 in Cre-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 of body weight, intraperitoneal) 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 scalp 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 1 s 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, Ecablemart.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.

16. Sohal, V.S., Zhang, F., Yizhar, O. & Deisseroth, K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* **459**, 698–702 (2009).