**Supplementary Methods**

**Mice**

58 TH::IRES-Cre C57BL/6J male mice expressing Cre recombinase under the control of the tyrosine hydroxylase promoter (TH::IRES-Cre), 14 TH::IRES-Cre C57BL/6J female mice and 6 DAT::IREs-Cre C57BL/6J male mice, all of ages 12-20 weeks, were used for these experiments. Mice were group-housed with 3-5 mice/cage on a 12-hr on, 12 hr off light schedule. All behavioral testing was performed during the light off time. All experimental procedures and animal care was performed in accordance with the guidelines set by the National Institutes of Health and were approved by the Princeton University Institutional Animal Care and Use Committee.

**Instrumental Reversal Learning Task**

Mice undergoing behavioral testing were water-restricted to 80-85% of their *ad-libitum* weight. All behavioral experiments were performed in a 21 x 18 cm modular operant chamber (MED Associates, ENV-307W). The nose port, levers and rewards were confined to one wall of the chamber. The beginning of each trial was signaled by the illumination of a central nose poke. Upon entering the nose poke, animals initiated the release of two levers spaced equidistantly right and left of the center nose poke after a random jittered delay between 0-1s. Animals were given 10s to press a lever, failure to do so resulted in an abandoned trial (<2% of trials for all mice) which was excluded from analysis. A lever press resulted in either the CS-
(0.5 s of white noise), and no reward delivery, or the CS+ (0.5 s of 5 kHz pure tone) and the delivery of 4ul reward (10% sucrose in H2O) in a custom reward port located central to the two levers and directly below the central nose poke. The time from lever press to CS also had a delay between 0-1s in 0.1s intervals sampled from a uniform distribution. All trials were separated by a constant 3s intertrial interval. Mice were trained daily and each behavioral session lasted one hour. Prior to data collection, mice were trained for 1-3 days to complete a nose poke and then a lever press in order to receive a reward (at this stage reward was not contingent on which lever was pressed). After 50 correct trials, the mice were then trained for 3-5 days on a paradigm where one lever was rewarded with 100% probability and the other lever was not rewarded, and which lever was rewarded was switched after at least 10 rewarded trials. Mice were then trained on the final paradigm for an additional 2 days before data was collected. Each block consisted of 10 correct trials plus a random number of trials drawn from a geometric distribution with a mean of 11 trials.

**NpHR photocurrent measurements in TH+ neurons**

Coronal slices containing the VTA/SNc were prepared from 3-4 month-old female TH:Cre mice (4 weeks after DIO-NpHR-EYFP virus injection). Mice were deeply anesthetized with isoflurane and decapitated to remove the brain. The isolated whole brains were immersed in ice-cold carbogenated NMDG ACSF solution (92 mM NMDG, 2.5 mM KCl, 1.25 mM NaH2PO4, 30 mM NaHCO3, 20 mM HEPES, 25 mM glucose, 2 mM thiourea, 5 mM Na-ascorbate, 3 mM Na-pyruvate, 0.5 mM CaCl2·4H2O, 10 mM MgSO4·7H2O, and 12 mM N-Acetyl-L-cysteine, pH adjusted to 7.3–7.4). Afterwards brain slices (300um) were sectioned using a vibratome (VT1200s, Leica, Germany) and then incubated in NMDG ACSF at 34°C for 15 minutes. Slices were then transferred into a holding solution of HEPES ACSF (92 mM NaCl, 2.5 mM KCl, 1.25 mM NaH2PO4, 30 mM NaHCO3, 20 mM HEPES, 25 mM glucose, 2 mM Thiourea, 5 mM Na-ascorbate, 3 mM Na-pyruvate, 2 mM CaCl2·4H2O, 2 mM MgSO4·7H2O and 12 mM N-Acetyl-L-cysteine, bubbled at room temperature with 95% O2/5% CO2). Whole-cell recordings were performed using a Multiclamp 700B (Molecular Devices, Sunnyvale, CA) using pipettes with a resistance of 3-5 MOhm filled with a potassium-based internal solution containing 120 mM potassium gluconate, 0.2 mM EGTA, 10 mM HEPES, 5 mM NaCl, 1 mM MgCl2, 2 mM Mg-ATP and 0.3 mM Na-GTP, with the pH adjusted to 7.2 with KOH. TH+ neurons were identified for recordings based on YFP expression. Photostimulation parameters were 560 nm and 5mW/mm2. Neurons were held at -70mV during photocurrent measurements. To confirm the ability of photocurrents to eliminate action potentials (**Supplementary Fig. 1c**), action potentials were induced by positive current injection (120pA, 100ms pulse duration, 5Hz).

**Optogenetic inhibition of putative dopamine neurons**

16 TH::ires-Cre male mice were bilaterally injected in the VTA (M/L +/- 0.5, A/P -3.1 and D/V -4.75 mm) with either AAV5 Ef1a-DIO-eNpHR3.0-EYFP (n=8 NpHR mice, UNC virus core, injected 0.5ul/ hemisphere of titer of 4 x 10^12 pp/ml) or AAV5-Ef1a-DIO-EYFP (n=8 control mice, UNC virus core, final titer of 4 x 10^12) virus, with littersmates randomly assigned to each group. 6 DAT::Cre male mice were bilaterally injected with AAV5 Ef1a-DIO-eNpHR3.0-EYFP. Optical fibers (300 um core diameter, .37NA) delivering 3-4 mW of 532 nm laser light (measured at the fiber tip) were implanted bilaterally above the VTA (M/L +/-0.4, A/P -3.3 and D/V -4.0
mm). Animals were anesthetized for implant surgeries with isoflurane (3-4% induction and 1-2% maintenance). Animals were given 5 days post-surgical procedure of recovery prior to behavioral testing. During behavioral sessions, light was administered randomly on 10% of trials beginning at the illumination of the central nose poke and remaining until either the duration of the CS- or the end or reward consumption. Days with optical stimulation alternated with days without stimulation; days without stimulation were excluded from analysis. All mice were tested without knowledge of the group in which the mouse belonged to known (in other words, the experimenter was blind). All segregation into groups performed after data collection. Anatomical targeting was confirmed as successful in all mice through histology after the experiment (Fig. 2b), and therefore no mice were excluded from this dataset.

Model of choice behavior

The mouse’s choice on trial \(i\) was modeled based on the outcomes and choices on the previous \(n\) trials with a logistic regression model:

\[
\log \left( \frac{C(i)}{1 - C(i)} \right) = \beta_0 + \sum_{j=1}^{n} \beta_j^R R(i-j) + \sum_{j=1}^{n} \beta_j^N N(i-j) + \text{error}
\]

where \(C(i)\) is the probability of choosing the right lever on trial \(i\). The rewarded press variable, \(R(i)\), was defined as +1 when the \(i\)th trial was both rewarded and a right press, -1 when the \(i\)th trial was rewarded and a left press and 0 when it was unrewarded. Similarly the unrewarded press variable, \(N(i)\), was defined as +1 when the \(i\)th trial was both unrewarded and a right press, -1 when the \(i\)th trial was unrewarded and a left press and 0 when it was rewarded. Together, these two variables define reward and choice history without redundancy. We chose these variables because of the ease of interpretation of the corresponding coefficients (\(j \text{Rand} j \text{N}\)), which reflect the strength of the relationship of rewarded (or unrewarded) outcomes in previous trials with the animal's current choice. For each mouse, the regression coefficients were fit using a logistic regression function in MATLAB (function \textit{glmfit}). For Figs. 1c and 2d, each coefficient point represents the average across mice\(^52\).

To examine the effect of optical inhibition of DA neurons on the animal's choice, we expanded the model to include the time of optical stimulation, as well as the interaction between choice and stimulation:

\[
\log \left( \frac{C(i)}{1 - C(i)} \right) = \beta_0 + \sum_{j=1}^{n} \beta_j^R R(i-j) + \sum_{j=1}^{n} \beta_j^N N(i-j) + \beta_0 + \sum_{j=1}^{n} \beta_j^{LR} L(i-j)R(i-j) + \sum_{j=1}^{n} \beta_j^{LN} L(i-j)N(i-j) + \beta_0 + \sum_{j=1}^{n} \beta_j^L L(i-j) + \text{error}
\]

where \(L(i)\) represents optical stimulation on the \(i\)th trial (1 for optical stim, 0 for control trials).

Finally, to statistically test the interaction between an animal’s choice, optical stimulation and the presence or absence of a light-activated opsin (i.e. NpHR vs YFP-control groups), we further augmented the model to fully cross all effects with an indicator for opsin. We estimated this model as a mixed effects model using the \texttt{lme4} package\(^51\) in the R statistical language (R...
Core Team, 2014) where the factors with repeated measures (intercept, stimulation, reward, and their interactions) were allowed to vary across animals.

**Calcium recordings in dopaminergic terminals**

11 mice were injected bilaterally with AAV5-CAG-Flex-GCamp6f-WPRE-SV40 (UPenn virus core, final titer of $3.53 \times 10^{12}$ pp/ml) in VTA using the coordinates described above. Optical fibers (400um core diameter, .48NA) were implanted in either the DMS (M/L +/-1.5, A/P 0.74 and D/V -2.4 mm) or the NAc (M/L +/-1.25, A/P 1.68 and D/V -4.15 mm). Each animal was implanted with one optical fiber in each brain region on opposite hemispheres, alternating which fiber was placed on which hemisphere between mice.

To record neural activity with gCaMP6f in terminals of dopaminergic neurons, light from the excitation laser (488nm) was chopped at 400 Hz before entering the implanted optical fiber (square wave with 50% duty cycle) and emission light was collected through the same fiber using a photodetector (New Focus, Femtowatt Photoreceiver model 2151), similar to 53. The excitation light was first filtered through a 488 nm UV/VIS bandpass filter (Thor Labs, FL488-10) to reduce background noise in the system. Emission light was filtered through a 525 nm GFP clean up filter (Thor Labs, MF525-39) prior to reaching the photodetector. The signal from the photodetector was then processed with a lock-in amplifier (Signal Recovery, Model 7265) to measure signal time-locked to the 400 Hz chopped excitation laser. The fluorescence signal from the lock-in amplifier was digitized at 15Hz on a data acquisition board (USB-201, Measurement Computing) and stored.

Targeting was confirmed as successful in all but one fiber through histology after the experiment, and therefore that recording site was excluded from analysis. Another recording site was excluded because there were no discernable gCaMP6f transients. No other data was excluded.

Animals expressing gCaMP6f were ran on the probabilistic reversal learning task. For the data in Figures 4 & 5, 3755 rewarded trials and 4003 unrewarded trials were obtained across 11 VTA/SN::NAc animals, for VTA/SN::DMS recordings there were 4861 rewarded trials and 5247 unrewarded trials. For the data in Figure 6, VTA/SN::NAc recordings contained 3526 ipsilateral and 4232 contralateral trials from 11 recording sites, VTA/SN::DMS recordings contained 4822 ipsilateral and 5286 contralateral trials from 11 recording sites. For the cell body VTA/SN::DMS recordings from Figure 7 there were 2312 ipsilateral and 2626 contralateral trials from 7 recordings sites.

**Calcium recordings in VTA/SN cell bodies**

44 TH::IRES-Cre male mice were injected bilaterally with AAV5-CAG-Flex-GCamp6f-WPRE-SV40 (UPenn virus core, final titer of $3.53 \times 10^{12}$ pp/ml) in DMS (M/L +/-1.4, A/P 0.74, D/V -2.6 mm). Optical fibers (400um core diameter, .48NA) were implanted in the VTA/SN (M/L +/-0.9, A/P -3.3 and D/V -4.2 mm). gCaMP6f recording were performed as described above and targeting was confirmed and reported in Fig. 7b. One data point was excluded because there were no discernable gCaMP6f transients.

**Analysis of calcium data**
Data were post-processed in MATLAB using a high-pass FIR filter with a passband of 0.4Hz, stopband of 0.1Hz and a stopband attenuation of 10dB to remove the baseline fluorescence and correct for drift in the baseline. dF/F was calculated by dividing the high-pass filtered signal by the mean of the signal prior to high-pass filtering. To compare signal across mice, the Z-score of dF/F was calculated by dividing each animal's dF/F trace by its standard deviation.

In order to calculate kernels that correspond to the isolated response to each behavioral event (Figs. 4e, 6c, 6f & 7f), the time-dependent gCaMP6f signal was modeled as the sum of the response to each behavioral event. The response to each behavioral event was considered as the convolution of a time series representing the time of the event as a series of 0’s and 1’s, where the 1’s correspond to the time of the event and the kernel corresponds to the response profile to that event. The model can be written as follows, where \( g(t) \) is the gCaMP6f signal, \( a, b, \) etc are example behavioral event time series (eg nose poke, lever presentation) and \( k_a \) and \( k_b \) are the kernels for the corresponding event:

\[
g(t) = g_0 + \sum_{t'=-1s}^{2s} a(t-t')k_a(t') + \sum_{t'=-1s}^{2s} b(t-t')k_b(t') + \cdots + \text{error}
\]

For each recording site, the coefficients of the kernels were solved using the method of least squares in MATLAB (with the function \textit{regress}). Figs. 4e 6c, 6f & 7f display the average (and SEM) kernel across recording sites, with each kernel based on data from 1 hr recording sessions across 4 days from one site. The calculation described in this paragraph is depicted schematically in Fig. 4a. This linear regression model successfully captured a significant amount of variance in the gCaMP6f response (R² of 0.24+/−0.020 for VTA/SN::NAc terminals and 0.11+/−0.016 for VTA/SN::DMS terminals).

\textbf{Voltammetry}

3 TH::IRES-Cre mice previously injected bilaterally with AAV5 CAG-Flex-GCamp6f-WPRE-SV40 in VTA (as described above) were used for these experiments (n=9 recording sites). Mice were anesthetized with urethane (1.5 g/kg, i.p.) and placed in a stereotaxic frame. A bipolar stimulating electrode (Plastics 1 Inc., Roanoke, VA) was placed in the medial forebrain bundle (MFB; M/L +/-1.1, A/P -2.4 and D/V -5 mm from bregma). An Ag/AgCl reference electrode was placed contralateral to the stimulating electrode in the left or right forebrain and held in place with a stainless steel skull screw. During the experiment, carbon-fiber microelectrodes (90-110 μm length) that had been coupled to 400 micron core optical fibers were placed into the DMS and NAc core (M/L +/-1.7, A/P +1.68 mm from bregma and D/V -3 to -4.3 mm from brain surface).

Changes in DA concentration during electrical stimulation of the MFB were assessed using fast scan cyclic voltammetry (FSCV). The carbon-fiber and Ag/AgCl reference electrodes were connected to an external headstage. Voltammetric recordings were made every 100 msec by applying a triangular waveform (-0.4 to +1.3 V, 400 V/sec). Data were digitized and stored to a computer using software written in LabVIEW (National Instruments, Austin, TX).

For each animal, a fresh carbon-fiber microelectrode was initially inserted 2.5 mm into the brain above the region of interest, and the triangular waveform was applied at a frequency of
60 Hz for 20 min. Waveform application was then changed to 10 Hz, and the electrode was lowered to the recording site. Electrical stimulation was applied to the MFB at a frequency of 60 Hz and pulse numbers varying between 24, 12, 6, and 1 biphasic pulse (300 μA, 4 ms each phase). These stimulation parameters were repeated 5 times at each recording site.

DA signals from FSCV were identified as previously described (Roitman et al., 2004). For analyte identification, current during a voltammetric scan is plotted against applied potential to yield a cyclic voltammogram (the chemical signature of the analyte). Cyclic voltammetric data were analyzed on stimulation trials. A background signal from 1 voltammetric scan (100 msec time bin) before a stimulation was subtracted from the remainder of the scans to reveal changes in DA concentration (rather than absolute values).

Current was examined at the oxidation potential of DA to reveal changes in DA current versus time during stimulation trials (I vs. t). After the experiment, changes in current were converted to changes in DA concentration by post calibration. Carbon-fiber microelectrodes were calibrated by recording the in vitro response to injections of known DA concentrations ranging from 300 nM to 3 μM in aCSF. The calibration curve was generated using 10 different electrodes and gave a linear response across dopamine concentrations with an R² value of 0.84. The equation generated from the linear response was used to determine an electrode response of 32.2 nA/μM.

**Histology**

After completion of behavioral experiments, mice were perfused with 4% PFA in PBS, and then brains were removed and postfixed in 4% PFA for 24 additional hours before transferring to 30% sucrose in PBS. After post-fixing, 40 micron sections were made with an American Optical 860 microtome. For immunohistochemistry, slices were blocked in 3% normal donkey serum (NDS) in PBS with 0.25% Triton X-100 for 30 minutes. Sections were then incubated at 4°C overnight in polyclonal chicken anti-TH primary antibody (1:1000, Aves Labs, No. TH) and in monoclonal rabbit anti-GFP primary antibody (1:1000, Life Technologies, No. G10362). PBS washes were performed to remove primary antibody, and slices were then incubated overnight in AlexaFluor488-conjugated Donkey Anti-Rabbit IgG (1:1000 dilution, Jackson IMMUNOResearch, No. 711-545-152) and AlexaFluor647-conjugated Donkey Anti-Chicken IgG (1:1000 dilution, Jackson IMMUNOResearch, No. 703-545-155). Following PBS washes, slices were mounted in 1:2500 DAPI in Fluoromount-G. Whole slices were imaged with a Nikon Ti2000E epifluorescent microscope to determine optical fiber targeting while confocal images of virus expression in cell bodies and terminals were taken using a Leica TCS SP8.

For the retrograde tracing experiments, 200-400 nl of Cholera Toxin subunit-B conjugated to either the Alexa-555 or -647 Fluorophore (Life Technologies C-34776 & C-34778) was injected into the NAc or DMS, respectively, and 5-7 days post-injection, animals were perfused as described above.

**Statistics**

All paired and unpaired t-tests mentioned in the Results were performed using MATLAB, using either the *ttest* or *ttest2* function. In all cases t-tests were two-tailed. The 2-way repeated measures ANOVAs reported in results paragraph 7 were done in R. The logistic regressions used to generate Figs. 1c and 2d and the linear regression used to generate the response
kernels (Figs. 4e, 6c, 6f & 7f) were done in MATLAB using the *glmfit* function. The estimate the standard deviation for the penetrance and specificity of gCaMP6f in TH+ cells (Results paragraph 6) we used the standard deviation for a binomial distribution: \( \sqrt{np(1-p)} \).

For all t-tests in this paper, data distributions were assumed to be normal but this was not formally tested. No statistical methods were used to pre-determine sample sizes but our sample sizes were similar to those generally employed in the field. A Supplementary Methods checklist is available.

We performed additional analyses on data of Figs. 6 & 7 that were not included in the main text, but that supported the same conclusions as the statistics that were included in the main text. The tests and results are described in detail below. Given the potentially longer time course of VTA/SN::NAc vs VTA/SN::DMS terminal responses (compare Figs. 6c,f), we assessed whether re-analyzing the VTA/SN::NAc terminal responses using longer analysis time windows would reveal significant differences in ipsilateral and contralateral responses. First, we expanded the size of the time bins used in the original t-test (see shaded region in Fig. 6f) from 100- to 200-ms and still found no significant difference between ipsilateral and contralateral response kernels for either nose poke or lever presentation (paired t-test; all 200-ms time-bins from -1s to 2s; \( p>0.01 \)). Second, to determine if a significant difference existed outside of the range of the initial time window of -1 to 2 seconds (kernels in Figs. 4, 6 & 7), we expanded the bin-wise analysis of VTA/SN::NAc terminal kernels from -2 to 3 seconds and again found no significant difference between ipsilateral and contralateral responses (paired t-test; all time bins; \( p>0.01 \)).

We also performed an additional statistical test to compare contralateral vs ipsilateral responses in VTA/SN::NAc terminals that did not involve bin-wise comparisons (Fig. 6b,c). Towards this end, we directly compared the peak response of ipsilateral and contralateral time-locked responses or kernels (Fig. 6b,c) across animals. This analysis revealed no significant difference in the time-locked responses between ipsilateral and contralateral trials (paired t-test; nose poke, \( p=0.24 \); lever presentation, \( p=0.78 \)) nor the nose poke or lever presentation kernels (paired t-test; nose poke, \( p=0.063 \); lever presentation, \( p=0.76 \)).

For the data in Fig. 6f, the bin-wise t-test comparing ipsilateral and contralateral response kernels (indicated as significant by the shaded grey area) for VTA/SN::DMS terminal recordings produced a minimum \( p \)-value of 2.51e-4 for the nose poke kernels and 8.84e-4 for the lever presentation kernel. Similarly, the same bin-wise analysis on the gCaMP6f time-locked responses in Fig. 6e results in a minimum \( p \)-value of 1.31e-7 for the nose poke and 1.55e-6 for the lever presentation. For the data in Fig. 7f, the bin-wise t-test comparing ipsilateral and contralateral response kernels (indicated as significant by the shaded grey area) for VTA/SN::DMS terminal recordings produced \( p \)-values as low as 3.22e-4 for the nose poke kernels and 0.005 for the lever presentation kernel. Similarly, the same bin-wise analysis on the gCaMP6f time-locked responses in Fig. 7e results in \( p \)-values as low as 4.54e-5 for nose poke and 6.44e-5 for lever presentation.

We performed an additional statistical test to compare contralateral vs ipsilateral responses in VTA/SN::DMS cell bodies that did not involve bin-wise comparisons. Directly comparing the peak of the contralateral kernel with the peak of the ipsilateral kernel (from -1s to +2s event onset) results in \( p=7.88e-4 \) for nose poke and \( p=0.012 \) for lever presentation (two-tailed paired t-test).
Finally, in the Results section, we compare the maximum difference in the contralateral and ipsilateral responses between VTA/SN::NAc and VTA/SN::DMS terminals (unpaired t-test; nose poke, p=.01, lever presentation, p=.024; pg. 7, para 3). Given the similarity of the VTA/SN::DMS terminal and cell body data (Figs. 6f & 7f), we also merged the two VTA/SN::DMS groups to make a larger combined group, and again performed the same statistical comparison between the combined VTA/SN::DMS and the VTA/SN::NAc terminal data. As expected, the p-values were smaller in this case due to the larger sample size (p=9.202- for nose poke and 0.011 for lever presentation).

**Code accessibility**

Custom MATLAB code used in this study is available upon request.

**Methods-only References**

51. Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker. “Fitting Linear Mixed-Effects Models Using lme4.” *arXiv:1406.5823 [stat]*, June 23, 2014. http://arxiv.org/abs/1406.5823.

52. Friston, K. J., K. E. Stephan, T. E. Lund, A. Morcom, and S. Kiebel. “Mixed-Effects and fMRI Studies.” *NeuroImage* 24, no. 1 (January 1, 2005): 244–52. doi:10.1016/j.neuroimage.2004.08.055.

53. Gunaydin, Lisa A., Logan Grosenick, Joel C. Finkelstein, Isaac V. Kauvar, Lief E. Fenno, Avishek Adhikari, Stephan Lammel, et al. “Natural Neural Projection Dynamics Underlying Social Behavior.” *Cell* 157, no. 7 (June 19, 2014): 1535–51. doi:10.1016/j.cell.2014.05.017.