博士論文

Dopamine D1 and D2 receptors in the nucleus accumbens regulate generalized conditioning and discrimination learning

（側坐核ドーパミンD1・D2受容体による
汎化的条件づけ・弁別学習の制御）

澤田　健

TABLE OF CONTENTS

Summary

Dopamine D2 receptors (D2Rs) are densely expressed in the striatum and have been implicated in neuropsychiatric disorders such as schizophrenia[1,2]. It has recently been shown by ex vivo experiments that the D2Rs detect transient reduction in dopamine concentration (DA dips), which has been linked to negative prediction error, to induce long term potentiation (LTP) in D2R-expressing spiny projection neurons (D2-SPNs) [3–6]. However, the behavioral function of the DA dip detection system by D2Rs has been elusive.

Here, I show in mice that reward conditioning mediated by dopamine D1 receptors (D1Rs) in the nucleus accumbens (NAc) induces remarkable stimulus generalization, and discrimination learning refines the generalization, which accompanies DA dips. Bidirectional optogenetic manipulations of those DA dips under the monitoring reveal that the DA dips causally control discrimination learning. Furthermore, Ca2+/calmodulin-dependent protein kinase II (CaMKII) and Adenosine 2A receptors (A2ARs), the signals necessary for the LTP in D2-SPNs, are also required for the learning. Unexpectedly, extinction learning does not involve DA dips or D2-SPNs. Treatment with methamphetamine, which dysregulates DA dips, and a D2R antagonist impairs and facilitates discrimination learning, respectively. These results suggest that

DA dips control discrimination learning via D2Rs to refine the generalized learning mediated by D1Rs.

Introduction

Animals survive in severe environments by learning from their experiences and adapting their behavioral strategies accordingly. One simple, yet important form of such adaptation is associative learning, whereby they learn the link between environmental stimuli and favorable outcomes. Such associative learning is thought to be mediated by prediction errors: difference between predicted and actual outcomes[7].

Many studies demonstrate dopamine neurons change their activities according to reward prediction error: they are tonically firing (4-8 Hz)[8], transiently activated up to 20~30 Hz (DA bursts) by more reward than predicted (positive prediction error), and transiently depressed (DA dips) by less reward than predicted (negative prediction error)[9–15]. This correlative link predicts the functional role of these phasic changes in dopamine neurons' activity as teaching signals to functionally drive associative learning[16].

The dopamine neurons mainly project to the striatum, including the nucleus accumbens (NAc)[17,18], where changes in firing rate lead to bidirectional transient changes in DA concentrations[19,20]. In the NAc, the primary cell type is the spiny projection neurons (SPNs), which are divided into two subpopulations, D1-SPNs and

D2-SPNs, according to almost dichotomous expression of dopamine D1 and D2 receptors (D1Rs and D2Rs)[21].

It has been well established that DA bursts drive long-term potentiation (LTP) of synapses in D1-SPNs via D1Rs[22] and reward associative learning[23–27]. By contrast, a recent study using ex vivo brain slices has demonstrated that DA dips are detected by D2Rs to disinhibit LTP in D2-SPNs[3]. In behavioral studies, optogenetic inhibition of dopamine neurons was sufficient to induce aversive learning[4,6]. However, how and for what endogenous DA dips are necessary has been unclear.

Here I established a method to optogenetically manipulate endogenous DA dips with simultaneous monitoring of DA neurons' axonal activities by fiber photometry. Using this method, I investigated the causal role of endogenous DA dips in two classical conditioning paradigms which induce negative prediction error: discrimination, where a cue is followed by reward while another is not despite generalized expectation, and extinction, where a cue once associated with reward comes to no longer predict reward.

The results show DA dips drive discrimination of stimuli which have been generalized in D1Rs-dependent reward conditioning. Drug infusion experiments and peptides expression experiments show that the LTP-related signals in D2-SPNs are also

required for the discrimination learning. Unexpectedly, extinction learning required neither DA dips nor D2-SPNs' LTP-related signals in the NAc. Treatment with methamphetamine, which dysregulated dopamine signals, and a D2R antagonist impaired and facilitated discrimination learning, respectively. These data suggest that endogenous DA dips, via the detection system by D2Rs, refine the generalized reward learning mediated by D1Rs.

Materials and Methods

Animals

Male C57BL/6J (Sankyo lab service), heterozygous DAT-IRES-Cre (B6.SJL-Slc6a3tm1.1 (cre) Bkmn/J, The Jackson Laboratory) and heterozygous A2A-Cre (B6.FVB (Cg)-Tg (Adora2a-cre) KG139Gsat/Mmucd) were used. Mice were housed in a 12 hour light/dark cycle with food and water available ad libitum except when water restriction was needed. The experimental protocol was approved by the Animal Experimental Committee of the Faculty of Medicine at the University of Tokyo. Sample size calculation was not performed and sample sizes were determined based on our previous study[25]. Group allocation was counterbalanced so that the equal number of experiments across groups were performed on the same day or the same periods. Experimenters were not blinded to the group allocation and the experiments were not randomized.

Plasmid constructions and adeno-associated virus preparation

For adeno-associated virus (AAV) production, I prepared the following constructs using PCR or Infusion (Takara):

pAAV-EF1-DIO-NLS-mCherry-W

pAAV-CaMKII (0.3)-DIO-NLS-mCherry-P2A-AIP-W

pAAV-EF1a-DIO-CsChrimsonR-mCherry-W

pAAV-EF1a-DIO-eNpHR3.0-mCherry-W,

pAAV-EF1-DIO-GCaMP6f-W.

Specifically, NLS (nuclear localization sequence; PKKKRKV), AIP (autocamtide 2-related inhibitory peptide (KKALRRQEAVDAL), a self-cleaving P2A peptide and the 74 N-terminal amino acids of CsChR for CsChrimsonR[28] were synthesized. I cloned ChrimsonR (#59171, Addgene), eNpHR3.0 (#26966, Addgene), and GCaMP6f (#40755, Addgene) from the plasmids. The pAAV-hSyn-GRAB$_{DA1m}$-W plasmid was a gift from Y. Li.

AAVs were produced as described previously[29], by co-transfection of pHelper (Stratagene), RepCap5 (Applied Viromics) and pAAV ITR-expression vectors above into HEK293 cells (AAV293, Stratagene). After three days, cells were harvested and AAVs were purified twice using iodixanol. Titers for AAVs were estimated using quantitative PCR.

Surgery

For head-restrained behavioural experiments, 6- to 7-week-old mice were anaesthetized with isoflurane and fixed in a stereotaxic instrument to install a head-plate with a donut-shaped ring, which could enclose drug infusion cannulas or optical fibres. The skull was exposed by removing the skin and glued to the head-plate with dental cement (RelyX Unicem2 clicker, 3M).

For the experiments involving drug infusion, a double guide cannula (26-gauge, 5.0 mm in length,1.5 mm apart, Plastic One) covered by dummy cannulas and dust caps (Plastic One) was implanted into the bilateral NAc (AP +1.25 mm, ML ±0.75 mm, DV +4.2 mm), and secured to the skull with dental cement.

For the experiments involving fibre photometry of GCaMP6f, 0.7 μl of AAV2/5-EF1-DIO-GCaMP6f (1.2 × 1013 genome copies (GC)/ml) was infused into the two sites of VTA (AP −3.1 mm, ML +0.55 mm, DV +4.2/4.55 mm) at 0.25 μl/min in DAT-Cre mice. After the AAV injection, a bundle of four optic fibres (CFML12U, ø200 μm, NA0.39, Thorlabs) was inserted at four sites in the NAc (anterior medial, AP +1.5 mm, ML +0.65 mm, DV +4.4 mm; anterior lateral, AP +1.5 mm, ML +1.35 mm, DV +4.4 mm; posterior medial, AP +1.0 mm, ML +0.65 mm, DV +4.4 mm; posterior lateral, AP +1.0 mm, ML +1.35 mm, DV +4.4 mm) using a scaffold holder made with 3D printing (DMM 3D print, Japan).

For experiments involving CsChrimsonR or eNpHR3.0, I infused 0.6 μl of a mixture of AAV2/5-EF1-DIO-GCaMP6f (1.2×10^{13} GC/ml) and AAV2/5-EF1a-DIO-CsChrimsonR-mCherry (1.2×10^{13} GC/ml), AAV2/5-EF1a-DIO-eNpHR3.0-mCherry (1.2×10^{13} GC/ml), or AAV2/5-EF1a-DIO-mCherry (1.2×10^{13} GC/ml) into each of the four sites in VTA (AP −3.1 mm, ML ±0.55 mm, DV +4.2/4.55 mm) of DAT-Cre mice. Immediately after the AAV injection, optic fibres (CFML12U, ø200 μm, NA0.39, Thorlabs) were implanted into the bilateral VTA (AP −3.1 mm, ML ±0.55 mm, DV +4.1 mm) and the right posterior lateral NAc (AP +1.0 mm, ML +1.35 mm, DV +4.4 mm).

For experiments involving fibre photometry of $GRAB_{DA1m}$ [30], 0.8 μl of AAV2/5-hSyn-$GRAB_{DA1m}$ (2.5×10^{13} GC/ml) was infused into the right posterior lateral NAc (AP +1.0 mm, ML +1.3 mm, DV +4.4 mm) at 0.5 μl/min in wild-type mice. Immediately after the AAV injection, optic fibres were implanted into the right posterior lateral NAc (AP +1.0 mm, ML +1.35 mm, DV +4.4 mm).

For experiments with D2-AIP, 1.5 μl of AAV2/5-CaMKII(0.3)-DIO-NLS-mCherry-P2A-AIP (2.0×10^{13} GC/ml) or AAV2/5-EF1a-DIO-NLS-mCherry (2.0×10^{13} GC/ml) was infused bilaterally into the lateral NAc (AP +1.25 mm, ML ±1.25 mm, DV +4.5 mm) in A2A-Cre mice.

After the surgery, the mice with and without virus injections were allowed to recover in their home cages for 3 days and 2–3 weeks, respectively.

Classical conditioning

Following recovery from surgery, 7- to 11-week-old mice went through at least four days of water restriction (receiving ~1 ml of water per day) and habituation to the experimental setup as described previously[25]. Tone–reward classical conditioning experiments began when mice weighed 80–85% of their free-drinking weight. In the conditioning session on day 1 of the generalization–discrimination task, mice were head-restrained and presented with a 1-s tone (CS+, 6 kHz, 70 dB) followed by 2 μl of 5% sucrose in water (US) in the CS+ with US trials 180 times. The delivery of sucrose water was controlled by a syringe pump (Legato 111, KD scientific). The inter-trial interval was chosen as a random sample from a uniform distribution bounded by 14 s and 20 s. For the generalization test on day 2,6-, 10-, 16-, or 28-kHz tones were pseudo-randomly presented three times each (Fig. 2, 3; 28 kHz was omitted in Fig. 3). Then mice underwent discrimination sessions on days 2 and 3 during which I presented a 6-kHz tone (CS+) with the US 40 times or a 10- or 16-kHz tone (CS−) without the US 80 times (40 times for the 10-kHz tone and 40 times for the 16-kHz tone in a

pseudo-randomized order) on each day (Fig. 4b, c, 9d, 13c, 16b, 20a and 23a). In some experiments, mice underwent a discrimination session only on day 2 (Fig. 10e and 17b). On the next day of the last discrimination session, the CS+ and CS− were presented five and ten times, respectively, without the US for a discrimination test. Then in Fig. 11 mice underwent extinction sessions on days 4 and 5 during which the CS+ was presented 30 times without the US on each day. On day 5, five trials of the CS+ presentation were added as an extinction test. In conventional discrimination task (Fig. 1b, c and 14b), mice were presented with 60 trials of the CS+ (6-kHz tone) with the US and 60 trials of the CS− (10-kHz tone) without the US for 3 days and then the discrimination test was performed on day 4. In extinction experiments, the CS+ was presented 40 times (Fig. 12e, g) or 20 times (Fig. 15b) per day without the US for two days after the conditioning sessions. Extinction tests were performed on the next day.

Drug infusion

To inject drugs into the NAc, a double cannula (28-gauge, Plastic One) was connected to a syringe (1801RN, 10 µl, Hamilton) through tubes containing the drug solution and was inserted into the guide cannula before the behavioural session. Drug infusion was conducted before (17 nl/min, 0.5 µl in total) and during (21 nl/min, 1.0 µl

in total) the behavioral sessions, which was controlled by a microsyringe pump (Legato 111, KD scientific). The drugs were dissolved in ACSF to the concentrations as below. SCH22290 (Wako), 400 μM; SCH58261 (ab120439, Abcam), 50 μM; sulpiride (190-12061, Wako), 600 μM. The spread of the infused drug was estimated by infusing 1 mM Alexa 594 hydrazide (A10438, Thermo Fisher Scientific).

Photometry and optogenetic stimulation

For four-site GCaMP photometry recording (Fig. 5)[31], light from a fibre-coupled LED (470 nm, M470F3, Thorlabs) was collimated (F220SMA-A, Thorlabs) to pass through an excitation filter (Thorlabs, MF469-35), galvanometer mirror (GVS011/M, Thorlabs), and dichroic mirror (Di01-R405/488/594, Semrock), and then focused through an objective lens (NA 0.50, UPlanFLN20x, Olympus) onto a custom-made fibre array that bundled four optical fibres (NA 0.50, Thorlabs). Patch cord fibres were connected to optic cannulas implanted at the NAc with mating sleeves after the mice were head-restrained. The output power at the tip of the fibres was adjusted to 5 μW. Emitted fluorescence passing through the filter (MF525-39, Thorlabs) was detected by a multialkali photomultiplier tube (PMT, H10722-20, Hamamatsu), whose signals were filtered by an in-line 1-kHz low-pass filter (EF110, Thorlabs). The

galvanometer mirror was controlled so that the light focused on one of the arrayed fibres for 10 ms. The LED was irradiated for 2 ms at each fiber and the mean value of PMT signals (digitized at 5 Hz) during this 2-ms period (10 time points) was recorded at 20 Hz. Leakage of excitation light into the unfocused bundle was less than 0.3% of intensity compared to when the light was focused. Signals from anterior lateral NAc and posterior medial NAc (open circles in Fig. 5b) were unreliable, possibly owing to the anterior commissure, so they were excluded from analysis. One-site photometry of GCaMP (Fig. 7a, 8a, 9a, 10a and 12a) or $GRAB_{DA1m}$ (Fig. 21 and 23) were performed as described above except that the galvanometer mirror was focused onto a single fibre.

For photometry recording from freely moving mice in their home cage (Fig. 22), light from a 405-nm LED (M405F1, bandpass filter: FB410-10, Thorlabs) was merged with light from 470-nm LED (M470F3, bandpass filter: FB470-10, Thorlabs) using a dichroic mirror (DMLP425R, Thorlabs). Excitation light was reflected by a dichroic mirror (DMLP505R, Thorlabs) and travelled through an optic fibre (RJPSL4, Thorlabs). Emission light was collected by a PMT (H10722-20, Hamamatsu; bandpass filter: MF525-39, Thorlabs). Excitation and sampling with 405-nm and 470-nm LEDs were conducted with 2-ms pulses for each and alternated at 20 Hz. To avoid bleaching,

a 30-s resting period was introduced after every 30-s recording period. Before recording, mice were habituated to the attached fibre for 10 min.

CsChrimsonR was stimulated by two LEDs (595 nm, M595F2, Thorlabs) connected to fibre cannulas implanted in the bilateral VTA. Stimulation power was adjusted for each mouse so that 20-Hz stimulation train (4-ms pulse, 10 times) yielded a similar DA photometry signal to the response to US sucrose water. During the behavioral sessions, 10 pulses with 4-ms width at 5 Hz or 20 Hz were applied just after the offset of the CS. The pulse timing was set so as not to interfere with photometry sampling. The stimulation was confirmed by the licking operant task where one lick triggered 10 pulses of 20 Hz or 5 Hz stimulation during 200 s sessions.

For eNpHR3.0 stimulation, continuous illumination for 2 s with blanking for 4 ms around 2-ms photometry sampling was applied from the offset of CS presentation, followed by ramping off over 45 ms.

Histology

Mice were anaesthetized using isoflurane, perfused with 4% paraformaldehyde, and decapitated. Brains were coronally sectioned into 50 μm slices using a vibratome (VT-1000, Leica). For immunostaining, slices were first incubated in 5% normal goat

serum containing PBS-T (0.2% Triton-X in PBS) at room temperature for 1 h, and then incubated at 4 °C for 48 h with primary antibodies: rabbit anti-tyrosine hydroxylase antibody (1:400, AB152, Millipore), or a mouse anti-mCherry antibody (1:500, 632543, Clontech) in PBS-T containing 5% normal goat serum. After washing with PBS, slices were incubated for 2 h at room temperature with secondary antibodies: goat anti-rabbit IgG antibody conjugated with Alexa 594 (1:400, A-11012, Thermo Fisher Scientific), a goat anti-rabbit IgG with Alexa 633 (1:400, A-21070, Thermo Fisher Scientific) or a rabbit anti-GFP with Alexa 488 (1:500, A-21311, Thermo Fisher Scientific). For nuclear staining, TO-PRO-3 (1:2,000, T3605, Thermo Fisher Scientific) was incubated simultaneously with the secondary antibodies. The slices were mounted using anti-fading reagents (H-1000, Vectashield, Vector Laboratories) and were imaged by confocal microscopy (TCS−SP5, Leica). For AIP expression analyses, 5 x–y images (0.65 μm/pixel) were acquired at intervals of 7 μm on z-axis.

To confirm fibre placements, I used the atlas data from the Allen mouse brain atlas (https://mouse.brain-map.org/static/atlas). The boundary between the NAc core and shell from another atlas[32] was overlaid.

Methamphetamine administration

Methamphetamine hydrochloride (MAP; Dainippon Pharma) was dissolved in PBS at 0.2 mg/ml, and administered intraperitoneally at 1 or 4 mg/kg. For the sensitization experiment, locomotor activity was measured as the distance travelled in a rectangular arena (41 cm × 31 cm). The test was conducted after habituation to the arena for 30 min. The positions of the mice were recorded with a CMOS camera (DCC1545M, Thorlabs) at 3 frames per second.

Conditioned place preference

Animals were group-housed in a cage containing paper bedding (pepper clean, SLC) before CPP. A box (41 cm × 31 cm) was separated into two equal-sized chambers by a black acrylic wall with a 5 cm × 5 cm square hole at the base. One zone contained novel bedding (sunfreak, Charles River) with dot-patterned walls, and the other contained different novel bedding (lignocel, J. Rettenmaier & Söhne) with striped walls. On day 1, the time spent in each chamber was measured during free exploration of both chambers for 15 min (pre-conditioning test). The mice that spent more than 10 min in either zone were excluded from the following experiments. On days 2 and 3, mice received an intraperitoneal injection and stayed in their home cage for 30 min before they were confined to one chamber for 30 min. After 4–6 h, the same protocol was

repeated, but in the other chamber. Mice were assigned to receive administration of MAP 1 mg/kg, MAP 4 mg/kg, or saline in one chamber. In the other chamber, mice always received a saline injection. The injection order of target drug and saline on a given day was counterbalanced across subjects. On day 4, the same procedure as the day 1 was conducted as a post-conditioning test. The positions of the mice were recorded with a CMOS camera (DCC1545M, Thorlabs) at 5 Hz and analyzed with DeepLabCut [33].

Data analysis

All of the licking and photometry data were collected using a DAQ board (PCIe-6323, National instruments) controlled by custom written scripts in Python 3. Water delivery with syringe pumps and LED excitation were also controlle by the system.

Statistical analyses were performed using Excel (Microsoft) and Excel Statistics (SSRI). Data are presented as mean ± s.e.m.

For analyses of the behavioral licking data, I calculated a lick score as the average lick frequency during CS presentation minus the baseline lick frequency during 2 s period before CS presentation. The lick score was averaged across each trial bin

indicated as below unless otherwise noted: conditioning phase, 45 trials of CS+ (but 25 trials of CS+ only in Fig. 15b) ; conditioning test, generalization test, discrimination test, and extinction test, 5 trials of CS+ or 10 trials of CS−; discrimination phase, 10 trials of CS+ or 20 trials of CS−; extinction phase, 10 trials of CS+ in Fig. 12e, g or 5 trials of CS+ in Fig. 15b. Mice that showed a lick score less than 2.5 for the CS+ at the initial 5 trials on day 2 were excluded from further analysis. A generalization index was calculated as (lick score for CS−)/(lick score for CS+). The extent of discrimination learning was evaluated by Δgeneralization index as (generalization index at DT) − (generalization index at GT). The extent of extinction was evaluated by an extinction index as (lick score at ET)/(lick score at CT).

For photometry signals in the head-restrained experiments, ΔF/F was calculated by comparing the average signal in 2 s baseline periods before each CS presentation, with the signal at each point during the trial. z-scores were calculated using the distribution during the baseline periods in all trials for each day and each mouse [34]. The scores were then down-sampled to 4 Hz. To quantify the Ca2+ activity of DA terminals using GCaMP, z-scored photometry signals were integrated during specific periods indicated below: DA dip, 1.5–3 s after CS onset (Fig. 5c, e, 7b, c, 9b, c, 10c, d, 11f, h and 12c, d); CS response, 0–1 s after CS onset (Fig. 5c, d). For

quantification of DA concentration using $GRAB_{DA1m}$, I used the peak or trough of z-scored photometry signals during specific periods: DA dip, trough during 1.5–3 s after CS onset (Fig. 23d); US responses, peak during 1–3 s after CS onset (Fig. 23e); CS responses, peak during 0–1 s after CS onset (Fig. 23f, g).

For photometry signals in experiments on freely-moving mice, motion artifacts were corrected by 405 nm signal, as previously described[34]. Signals of 405 nm in the baseline period were fitted to signals of 470 nm by a polynomial (ax + b) equation, in which coefficients were determined by least-squares methods. Then, ΔF/F was calculated by ([470 nm signal] – [fitted 405 nm signal])/[fitted 405 nm signal].

For the analysis of fluorescent protein expression (Fig. 13b), original confocal images were down-sampled to 32.5 μm/pixel, z-stacked, and then registered across subjects. The average image was calculated after the autofluorescence signal value detected at the neocortex was subtracted from each image.

Results

DA dips during discrimination learning

As DA dips has been reported to occur when animals encounter the omission of an expected reward, or negative prediction error[16,17,35], I designed a discrimination task that involves negative prediction error. Water-deprived mice were head-restrained and then presented with a conditioned stimulus (CS+, 6-kHz tone for 1 s) with an unconditioned stimulus (US, 2 μl of sucrose water) immediately thereafter, and CS− (a 10-kHz tone) without a US, 60 times each pseudo-randomly for 3 days (Fig. 1a–e). During this discrimination learning task, conditioned licking responses to the CS− was initially comparable to those to the CS+ (Fig. 1c, d), suggesting that conditioning with the CS+ was generalized to the CS−[10,13,36]. On day 4, mice showed less conditioned licking responses to the CS− than to the CS+(Fig. 1c–e), indicating that the mice could discriminate between the reward-predicting CS+ tone and the non-reward-predicting CS− tone.

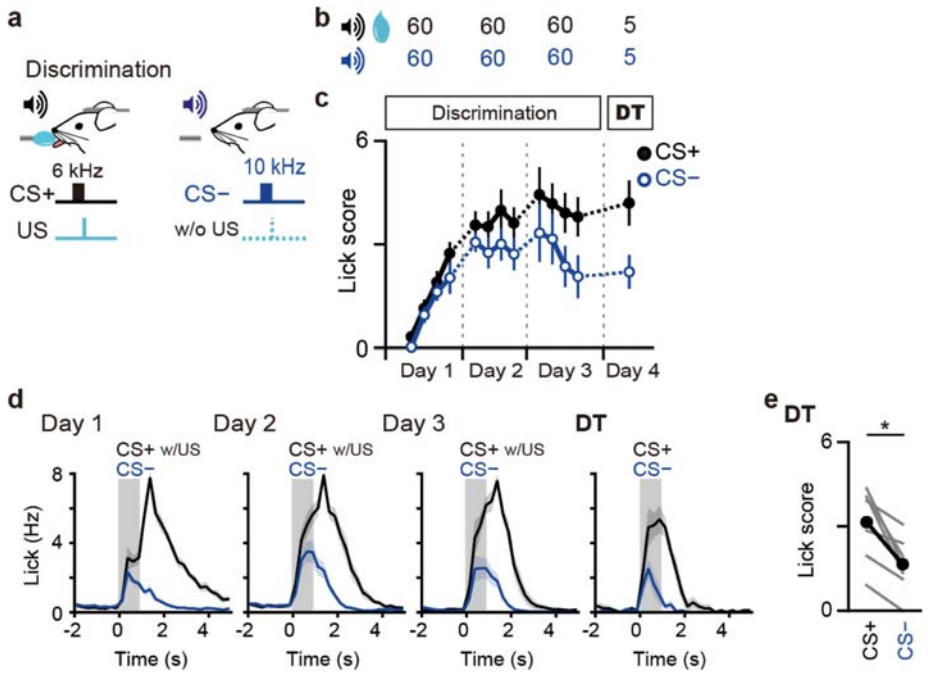

Figure 1. Discrimination learning task.

a, Schematic of the behavioural setup for the tone–reward discrimination task in classical conditioning in head-restrained mice. The tube monitored licking responses. **b**, Numbers of trials during the discrimination task. The order of presentation was pseudo-randomized. DT, discrimination test. **c**, Time course of lick scores aligned with numbers of trials in **b.** n = 7 mice. **d**, Peristimulus time histograms (PSTHs) of licking responses plotted against time from CS onset on days 1, 2, 3, and 4. The grey bar indicates the period of tone presentation. Vertical grey shading indicates CS periods. n = 7 mice **e**, Lick scores for CS+ and CS− on day 4. Two-sided Wilcoxon signed-rank test (n = 7 mice, *P = 0.016). Error bars and shading indicate s.e.m. The numbers of mice are biological replicates.

To separately investigate the simultaneous processes of generalization and discrimination, I first quantified the generalization of tone–reward conditioning after paring of the CS+ with the US 180 times on day 1 (Fig. 2). On day 2, the mice showed conditioned responses not only to the 6-kHz tone but also to tones up to 28 kHz, to which the mice had not been exposed during the conditioning on day 1 (Fig. 2a–d), although they did not show conditioned responses to light (Fig. 2e-g). These indicate that the reward learning is intrinsically generalized across a broad range within a stimulus modality. The generalized conditioned responses tested on day 2 were blocked by a D1R antagonist (SCH23390) injection into the NAc during conditioning on day 1 (Fig. 3).

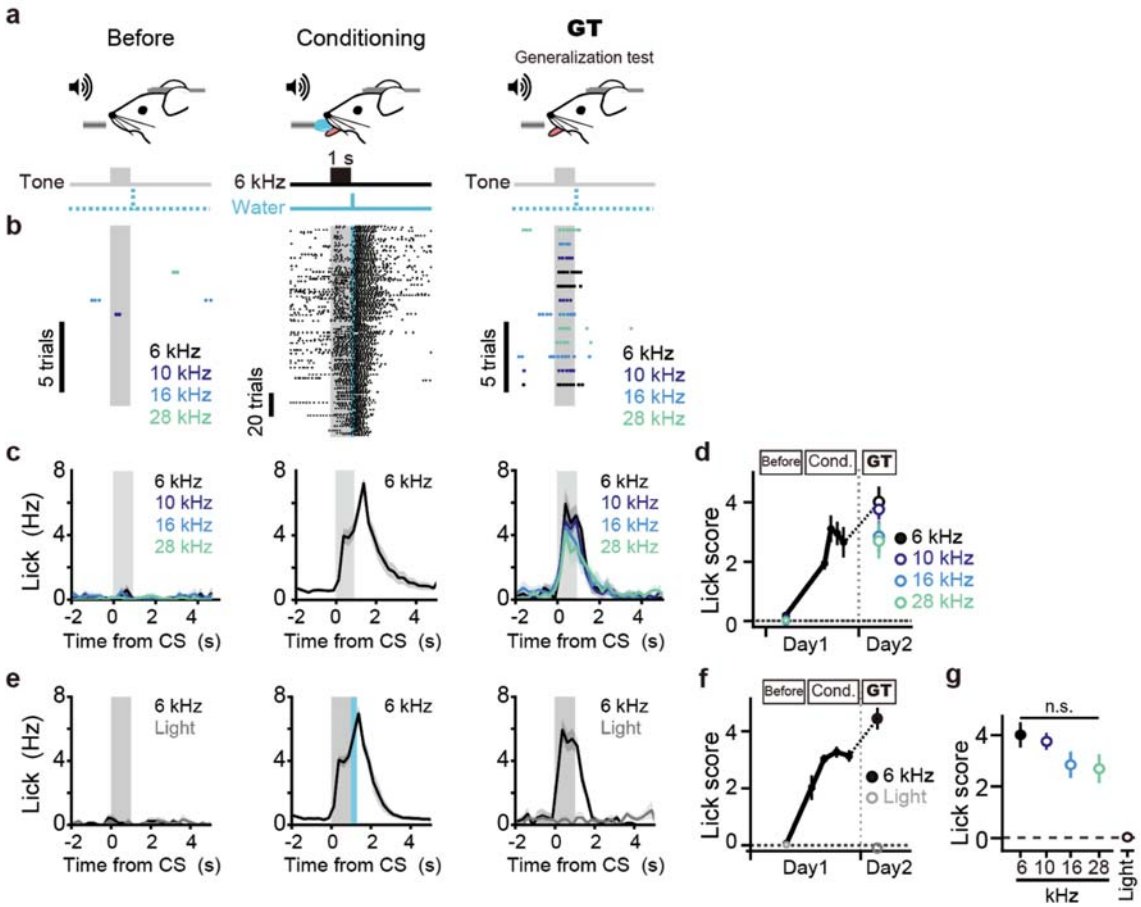

Figure 2. Generalization of cues after reward conditioning.

a, Schematic of the tone–reward classical conditioning task for head-restrained mice. **b**, Raster plots of licking responses before (left), during (middle) and after (right, GT, generalization test) conditioning for a representative mouse. Grey shading indicates the periods of CS presentation. **c**–**f**, Generalization to pure tones (**c**, **d**), and light (**e**, **f**). PSTHs in **c**, **e** show licking responses before (left), during (middle) and after (right) conditioning. Vertical grey shades indicate CS periods. **d**, **f** show lick scores are plotted against time. $n = 11$ (**c**, **d**) and 7 mice (**e**, **f**). **g,** Lick scores at generalization test. $P = 0.14$ by Friedman's test. Error bars and shade indicate s.e.m. The numbers of mice are biological replicates.

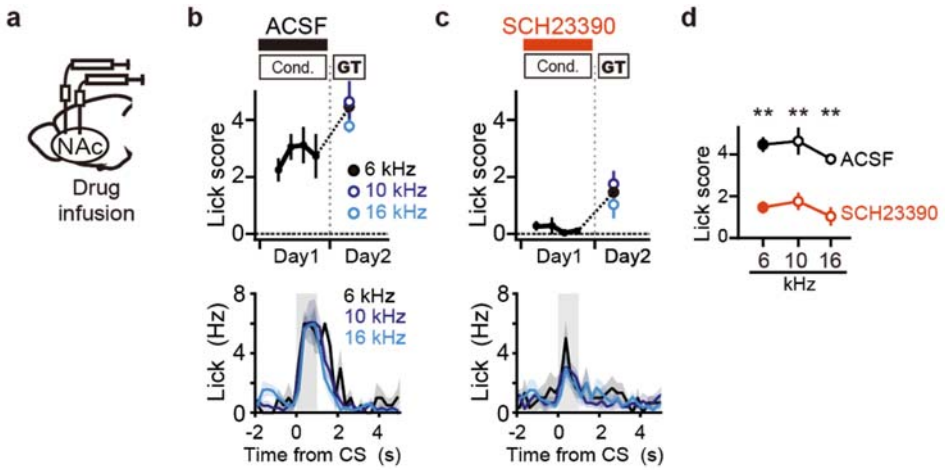

Figure 3. Effect of SCH23390 on the generalized conditioned responses.
a, Schematic of drug infusion into the bilateral NAc. **b**, **c**, Effect of drug infusion on generalized conditioned responses. Top, plots of lick scores; bottom, PSTHs at generalization test (GT). Grey shade indicates periods of CS presentation. Black and red bars indicate infusion periods. $n = 6$ (**b**) and 6 mice (**c**). Error bars and shades indicate s.e.m. **d**, Lick scores at generalization test. Two-sided Mann–Whitney U test with Bonferroni correction (6 kHz, $^{**}P = 0.002$; 10 kHz, $^{**}P = 0.009$; 16 kHz, $^{*}P = 0.002$). The numbers of mice are biological replicates.

I then tested whether mice could discriminate non-reward-predicting tones from the reward-predicting tone after generalized reward conditioning. I presented a 10- or 16-kHz tone (CS−) without a US 80 times on days 2 and 3, while maintaining the pairing of the 6-kHz tone (CS+) and US (Fig. 4a-c). Simultaneously, the activity of DA axons was monitored at the lateral and medial parts of the NAc core by fibre photometry of GCaMP6f, which was expressed in DA neurons in the ventral tegmental area (VTA) (Fig. 5a, b). After conditioning with the CS+ on day 1, mice showed generalized licking and DA responses to the CS+ and CS− at the generalization test on day 2 (Fig. 4c-e and 5c, d, GT). During the discrimination session on days 2 and 3, DA dips were preferentially observed in the lateral part of the NAc core at the timing of the reward omission in CS− trials (Fig. 5c, e, Discrimination). At the discrimination test on day 4 mice showed suppressed licking and DA responses to the CS− (Fig. 4c-f and 5c, d, DT). Mice could also perform other patterns of discrimination learning (Fig. 6). DA dips were also observed when mice performed the conventional discrimination task (Fig. 7a–c). Thus this generalization–discrimination paradigm (Fig. 4c) allowed us to evaluate the effects of reward omission on DA dips and discrimination learning separately from the acquisition of reward conditioning. The results suggest that mice

learned discrimination of the non-reward-predicting CS− from the reward-predicting CS+ during the task on days 2 and 3.

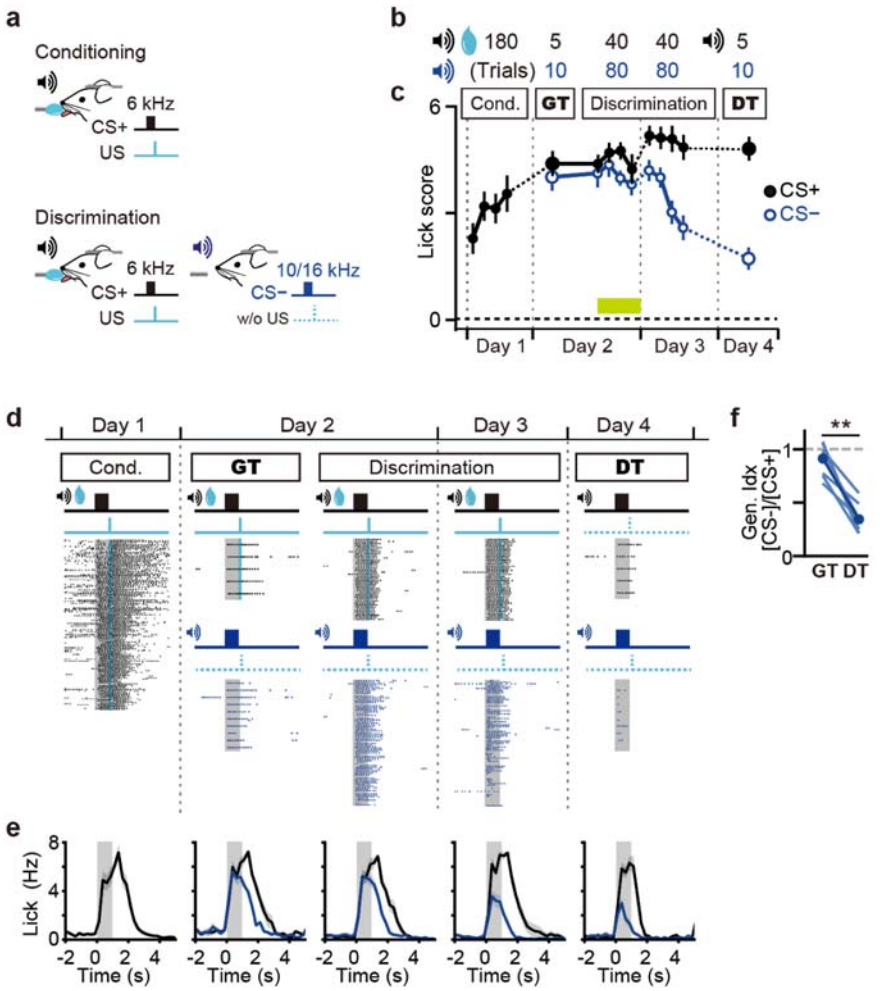

Figure 4. Behavioral task for generalization and discrimination

a, Schematic of the conditioning and discrimination task . CS− trials during discrimination task consisted of 40 trials at 10 kHz and 40 trials at 16 kHz. **b**, Numbers of trials during each session. **c**, Time course of lick scores**.** *n* = 9 mice. GT, generalization test; DT, discrimination test. **d**, **e**, Representative raster plots (**d**) and averaged PSTHs (**e**) of lick responses during generalization and discrimination. *n* = 9 mice. **f,** Generalization indices ((lick score in CS− trials) / (lick score in CS+ trials)) for trials shown in **c.** ***P* = 0.004 by two-sided Wilcoxon signed rank test. The numbers of mice are biological replicates.

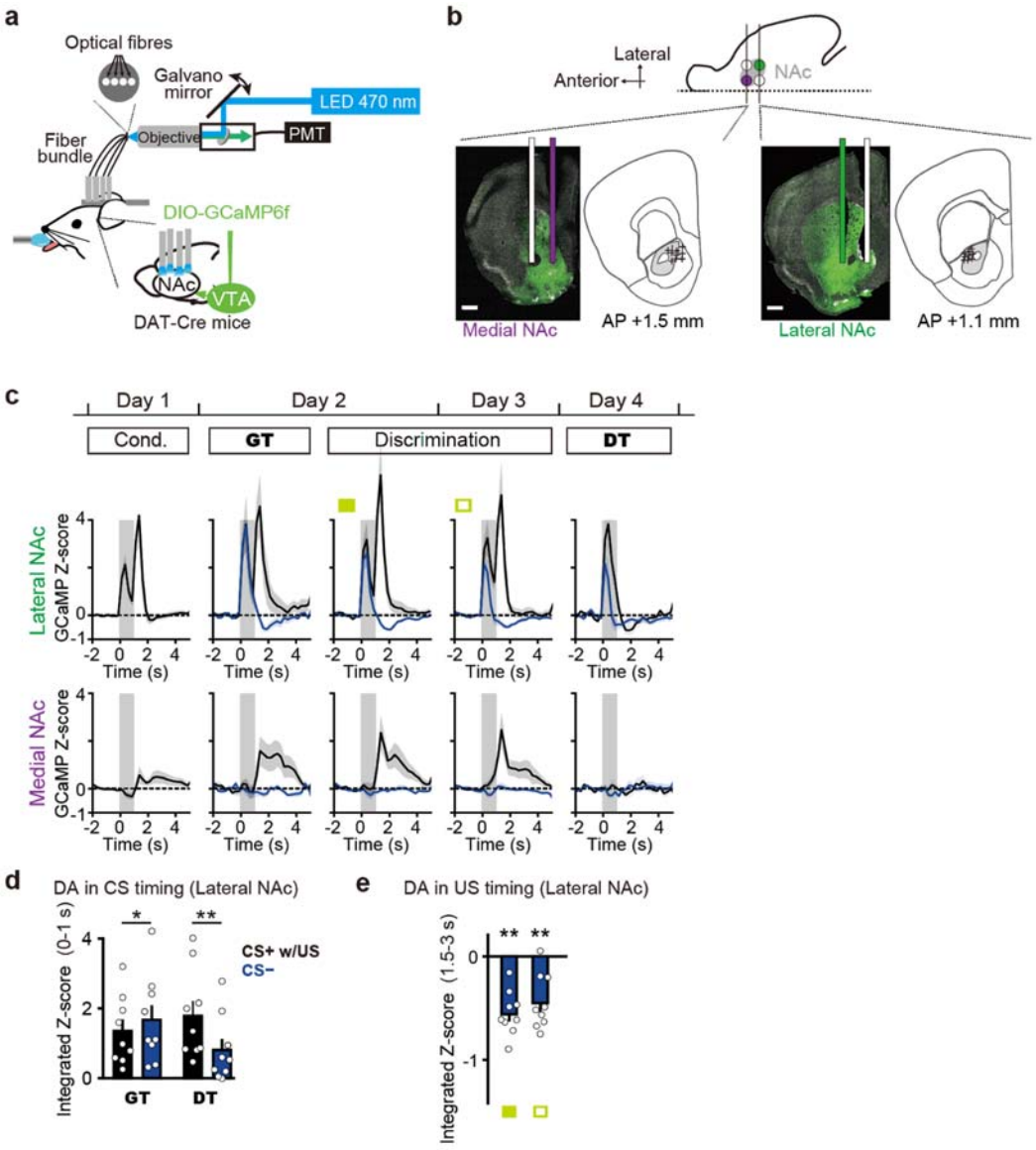

Figure 5. Photometry of DA terminal activity at sub-regions of the NAc during the generalizaion-discrimination task.

a, Schematic of four-site (anterior or posterior/medial or lateral) photometry. **b**, Representative confocal images showing locations of four fibres in the NAc. Signals were stably obtained from the medial NAc at anterior position (left) and the lateral NAc core at posterior position (right). The locations of the tips of fibres were plotted on the reference atlas. Grey areas indicate the NAc core. **c,** Averaged photometry signals obtained from the lateral (top) and medial NAc (bottom) during the task shown in Fig. 4. $n = 9$ mice. Vertical grey shades indicate CS periods. GT, generalization test; DT, discrimination test. **d**, Plots of photometry signals in the lateral NAc at CS timing (0–1 s after CS onset) for trials shown in Fig. 4c. Two-sided Wilcoxon signed-rank test ($n = 9$ mice, generalization test, $*P = 0.020$; discrimination test, $**P = 0.004$). **e**, Plots of photometry signals at US timing (1.5–3 s after CS onset). Green boxes indicate the

periods of sampling indicated in **c**. Two-sided Wilcoxon signed-rank test (vs baseline) with Bonferroni correction for multiple comparisons (n = 9 mice, day 2 discrimination, $^{**}P$ = 0.004; day 3 discrimination, $^{**}P$ = 0.008). Error bars and shading indicate s.e.m. The numbers of mice are biological replicates.

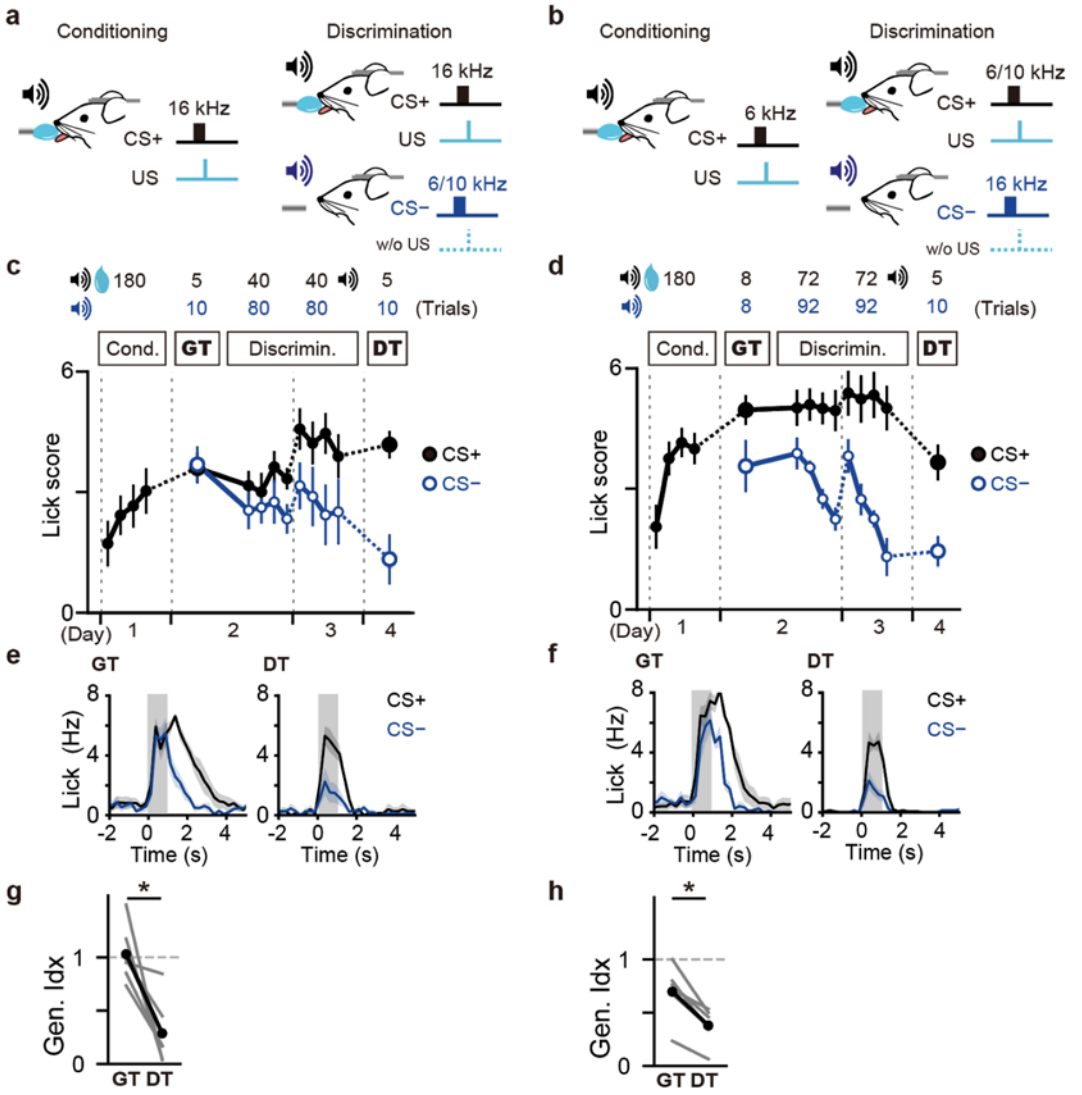

Figure 6. Generalization and discrimination learning of various tone combinations for CS+ and CS−.

a, **b**, Schematic for assignment of the tones to CS+ and CS− trials in variations of generalization–discrimination learning tasks. **c**, **d**, Time course of lick scores. Trial numbers for each trial type are indicated above the plots. In CS− trials in **c**, equal numbers of trials were pseudo-randomly allocated to 6-kHz or 10-kHz tone. For example, there were 40 trials at 6 kHz and 40 trials at 10 kHz during discrimination. In CS+ trials in **d**, only 6-kHz tones were presented during conditioning and discrimination test, and 6 kHz and 10 kHz were presented pseudo-randomly for 36 trials each during discrimination. $n = 6$ mice per group. GT, generalization test; DT, discrimination test. **e**, **f**, Averaged PSTHs of lick responses at generalization test (left, GT) and discrimination test (right, DT). $n = 6$ mice per group. **g**, **h**, Generalization indices at generalization test and discrimination test. Two-sided Wilcoxon signed-rank test (**g**, $n = 6$ mice, $^*P = 0.03$;

h, $n = 6$ mice, $*P = 0.03$). Error bars and shading indicate s.e.m. Vertical grey shading in **e**, **f** indicate CS periods. The numbers of mice are biological replicates.

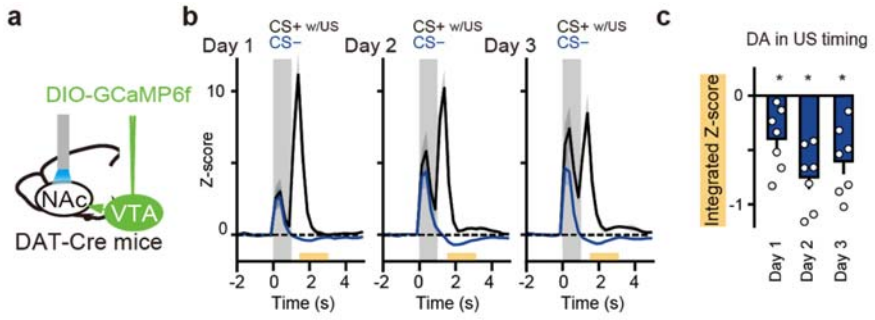

Figure 7. Photometry of DA terminal activity at the lateral NAc during the conventional discrimination task.

a, Schematic of fibre photometry of DA activities in the lateral NAc. **b**, Averaged *z*-scores of green fluorescence during the conventional discrimination task shown in Fig. 1. Orange bars indicate integration periods. Vertical grey shading indicates CS periods. $n = 7$ mice. **c**, Plots of photometry signals in US timing (1.5–3 s after CS onset). Two-sided Wilcoxon signed-rank test vs baseline with Bonferroni correction ($n = 7$ mice, day 1, $^*P = 0.016$; day 2, $^*P = 0.016$; day 3, $^*P = 0.016$). Error bars and shading indicate s.e.m. The numbers of mice are biological replicates.

Role of DA dips during discrimination learning

DA dips have been expected to act as negative prediction error signals[16], thus driving suppression of reward prediction. To address whether the observed DA dips had a causal role in discrimination learning, I optogentically stimulated DA neurons with CsChrimsonR expressed in the VTA, while monitoring DA neurons' axonal activity at the lateral NAc core using GCaMP (Fig. 8a, b and 9a-d). For each mouse, the stimulation power was adjusted so that DA photometry responses to phasic 20-Hz stimulation (595-nm LED, 4-ms pulse, 0.5 s, 0.5–1 mW) matches those to sucrose water delivery (Fig. 8c). With this intensity, 20-Hz DA stimulation reinforced licking behaviour in an operant task, indicating that the stimulation of DA worked effectively (Fig. 8d, e). By contrast, 5-Hz DA stimulation for 2 s only slightly increased the photometry signal and did not reinforce the licking behaviour in the operant task (Fig. 8c-e), consistent with in vivo baseline DA firing at around 4–8 Hz[37,38].

During the discrimination task, stimulation of DA neurons at 5 Hz with the adjusted power for 2 s after CS− offset ablated the DA dips (Fig. 9b, c). This optogenetic manipulation inhibited discrimination learning as shown in the discriminaion test on day 4 (Fig. 9d–f). By contrast, the inhibition of DA neurons with eNpHR3.0 for 2 s after CS− offset enhanced the DA dip during the discrimination task

(Fig. 10a-d) and facilitated one-day discrimination learning (Fig. 10e-g). These data indicate that DA dips bidirectionally contol discrimination learning.

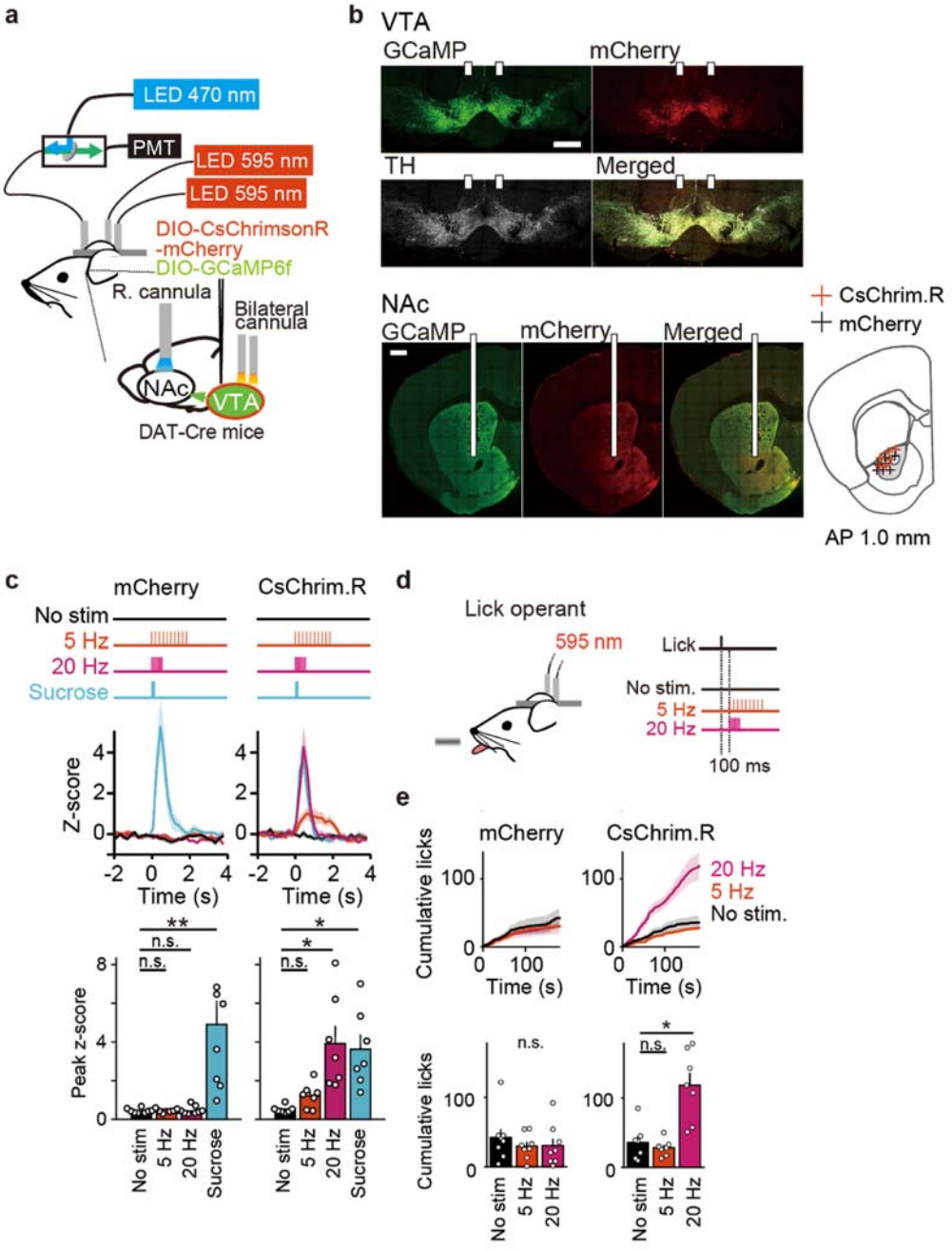

Figure 8. Methods for optogenetic manipulation of DA neurons under monitoring with fiber photometry.

a, Schematic of experimental setup for optogenetic manipulation of DA neurons in VTA under monitoring of axonal activities at NAc with fiber photometry. AAVs were injected to express CsChrimsonR and GCaMP6f in VTA-DA neurons. Optical fibres were implanted at the VTA for stimulation of CsChrimsonR and at the right NAc for photometry of GCaMP. **b**, Representative confocal images of virus expression and fibre position. TH, tyrosine hydroxylase immunostaining. White bars indicate traces of

optical fibres. The locations of the fibre tips were plotted on the reference atlas (lower right). Grey areas in the atlas indicate the NAc core. **c**, Photometry of DA neurons stimulated with red LED. Top, schematic for stimulus pattern; middle, photometry responses at NAc; bottom, plots of peak photometry responses. Friedman's test for mCherry (n = 8 mice, P = 0.001; post hoc two-sided Wilcoxon signed rank test with Bonferroni correction, no stimulation versus 5 Hz, P = 0.26; no stimulation versus 20 Hz, P = 0.89; no stimulation versus sucrose, $^{**}P$ = 0.008) and for CsChrimsonR (n = 7 mice, P = 0.0003; post hoc two-sided Wilcoxon signed rank test with Bonferroni correction, no stimulation versus 5 Hz, P = 0.093; no stimulation versus 20 Hz, $^{*}P$ = 0.016; no stimulation versus sucrose, $^{*}P$ = 0.016). **d**, Schematic of operant licking task. Licking behaviour triggered optogenetic self-stimulation of DA neurons in the VTA without water delivery. **e**, Cumulative lick number against time (top) and plots across the conditions (bottom). Friedman's test for mCherry (n = 8 mice, P = 0.60) and for CsChrimson (n = 7 mice, $^{**}P$ = 0.004; post hoc two-sided Wilcoxon signed rank test with Bonferroni correction, no stimulation versus 5 Hz, P = 0.31; no stimulation versus 20 Hz in CsChrimsonR, $^{*}P$ = 0.016). The numbers of mice are biological replicates.

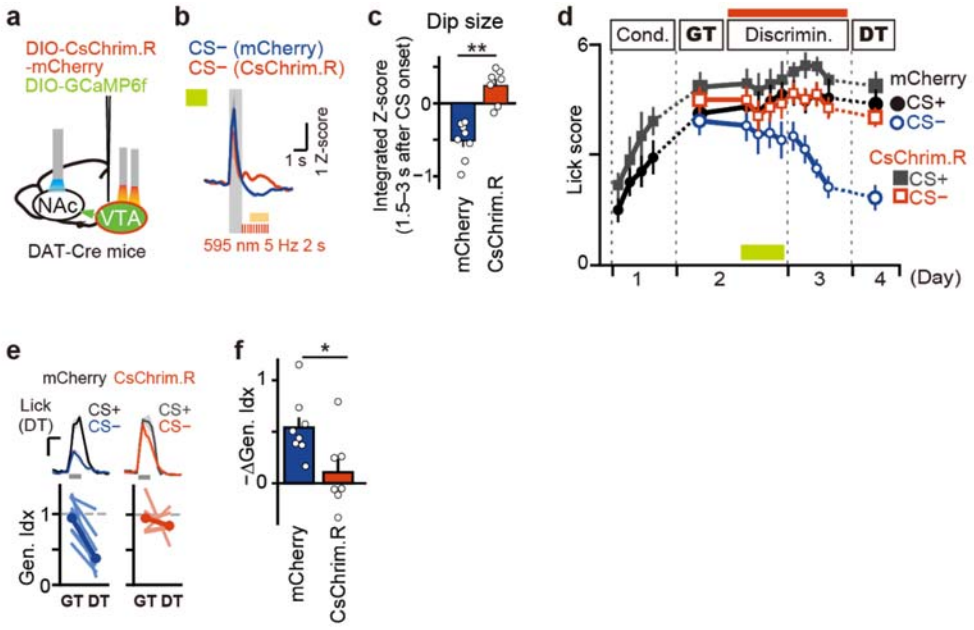

Figure 9. Optogenetic ablation of DA dips during discrimination learning.
a, Schematic of experimental setup for optogenetic ablation of DA dips under monitoring with fiber photometry during discrimination. **b**, Averaged *z*-scores of photometry responses in the NAc during the discrimination period indicated with the green bar in **d**. LED stimulation (red bars) was imposed for 1–3 s after CS onset in all CS− trials during the discrimination learning. **c**, Integrated *z*-score of photometry responses 1.5–3 s after CS onset (**b**, orange bar). $^{**}P = 0.0003$ by Two-sided Mann–Whitney U test. **d**, Time course of lick scores. Red bar indicates the period of optogenetic stimulation. GT, generalization test; DT, discrimination test. **e**, Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **f**, ΔGeneralization indices. $^{*}P = 0.029$ by Two-sided Mann–Whitney U test. $n = 8$ (mCherry) and 7 (CsChrimsonR) mice. The numbers of mice are biological replicates.

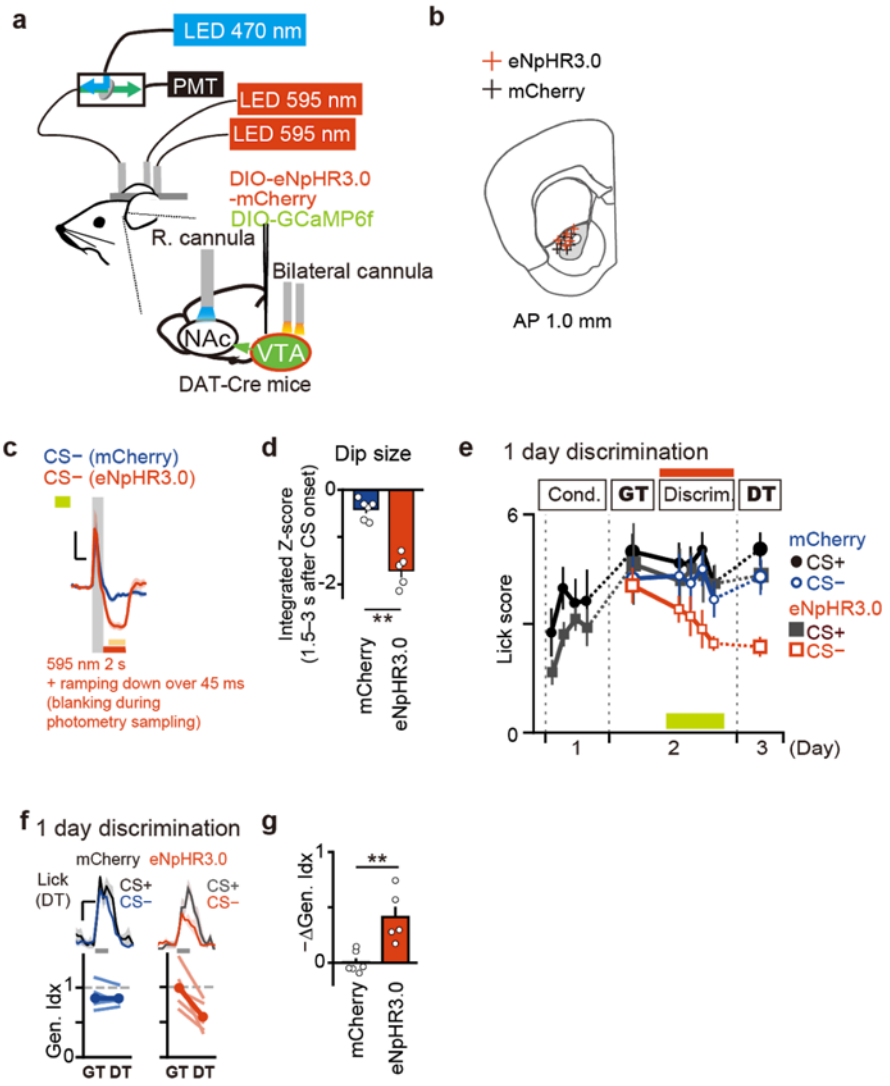

Figure 10. Optogenetic enhancement of DA dips during discrimination learning. **a**, Schematic of experimental setup for optogenetic ablation of DA dips under monitoring with fiber photometry during discrimination. AAVs were injected to express eNpHR3.0 and GCaMP6f in VTA-DA neurons. Optical fibres were implanted at the VTA for stimulation of eNpHR3.0 and at the right NAc for photometry of GCaMP. **b**, Locations of the fibre tips plotted on the reference atlas. **c**, Averaged *z*-scores of photometry responses during CS− trials in the discrimination period indicated with the green bar in **e**. **d,** Integrated *z*-score of photometry responses 1.5–3 s after CS onset (**c**, orange bar). ***P* = 0.004 by Two-sided Mann–Whitney U test. **e**, Time course of lick scores. Red bar indicates the period of optogenetic stimulation. GT, generalization test; DT, discrimination test. **f,** Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **g**, ΔGeneralization indices. ***P* = 0.004 by Two-sided Mann–Whitney U test. *n* = 6 (mCherry) and 5 (eNpHR3.0) mice. The numbers of mice are biological replicates.

Role of DA dips during extinction learning

I next investigated whether DA dips were generally induced by the omission of expected rewards, using an extinction paradigm. After the generalization–discrimination task, the CS+ was presented without the US on days 4 and 5, leading to gradual decrease of the licking response to the CS+. (Fig. 11a-e, g, Extinction). During this extinction task, DA dips were observed upon reward omission for the initial ten trials (Fig. 11f, h, Day4 Extinction, early phase). Unexpectedly, however, during the remaining period, DA dips were not significant at either of the lateral or medial NAc, despite remaining licking and DA responses upon the presentation of CS+ (Fig. 11f, h, Extinction). These results suggest that DA dips became weak when the mice experienced the absence of the US in a consecutive manner.

To further investigate the involvement of DA dips, I applied the 5-Hz DA stimulation in the VTA right after CS+ during the extinction task (which was conducted directly after the conditioning task), and found that this manipulation did not significantly inhibit extinction (Fig. 12a–e, h, i). By contrast, 20-Hz stimulation did inhibit extinction (Fig. 12f–i), as previously described[23]. These results suggest that the DA dip in the NAc core is a selective signal generated in a reward-predictable context

such as in discrimination learning, although I could not exclude the possibility that DA dips outside the NAc core regulate extinction, given the heterogeneity of DA neurons.

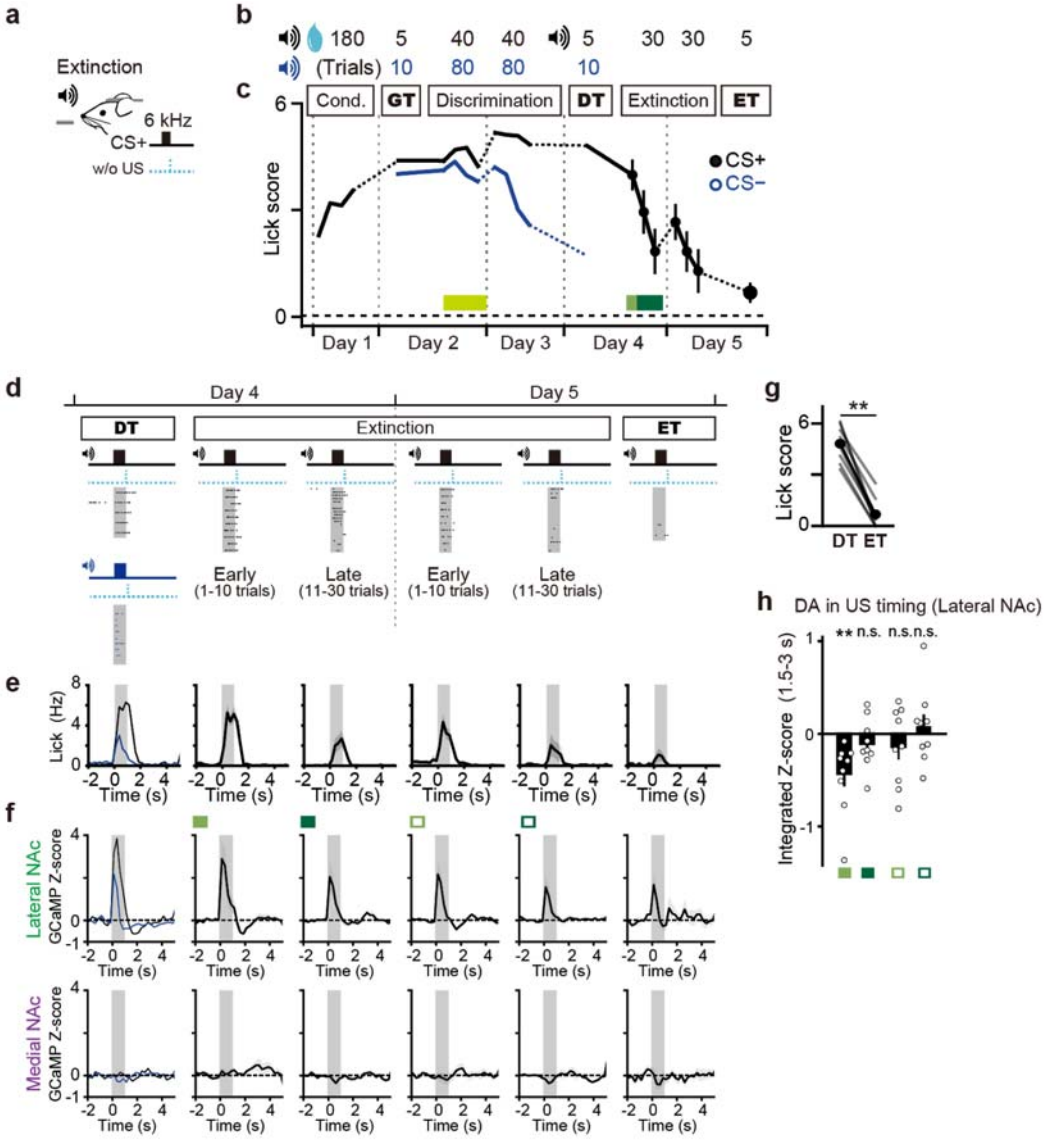

Figure 11. Behavioral and DA terminal responses in extinction task.

a, Schematic of extinction task. **b**, Numbers of trials during each session. **c**, Time course of lick scores during conditioning, discrimination and extinction task. $n = 9$ mice. GT, generalization test; DT, discrimination test; ET, extinction test. Lick scores during the period until discrimination test are from Fig. 4c. **d**, **e**, Representative raster plots (**d**) and averaged PSTHs (**e**) of lick responses during extinction. Those at discrimination test are from Fig. 4d, e. Vertical grey shades indicate CS periods. $n = 9$ mice. **f,** Averaged photometry signals obtained from the lateral (top) and medial NAc (bottom). Those at discrimination test are from Fig. 5c. **g,** Lick scores at the CS+ presentation in the trials shown in **c.** $^{**}P = 0.004$ by two-sided Wilcoxon signed rank test. **h**, Plots of photometry signals at US timing (1.5–3 s after CS onset). Green boxes indicate the periods of sampling indicated in **f**. Two-sided Wilcoxon signed-rank test (vs baseline) with

Bonferroni correction for multiple comparisons (n = 9 mice, day 4 extinction early, $^{**}P$ = 0.004; day 4 extinction late, P = 0.43; day 5 extinction early, P = 0.57; day 5 extinction late, P = 0.82). Error bars and shading indicate s.e.m. The numbers of mice are biological replicates.

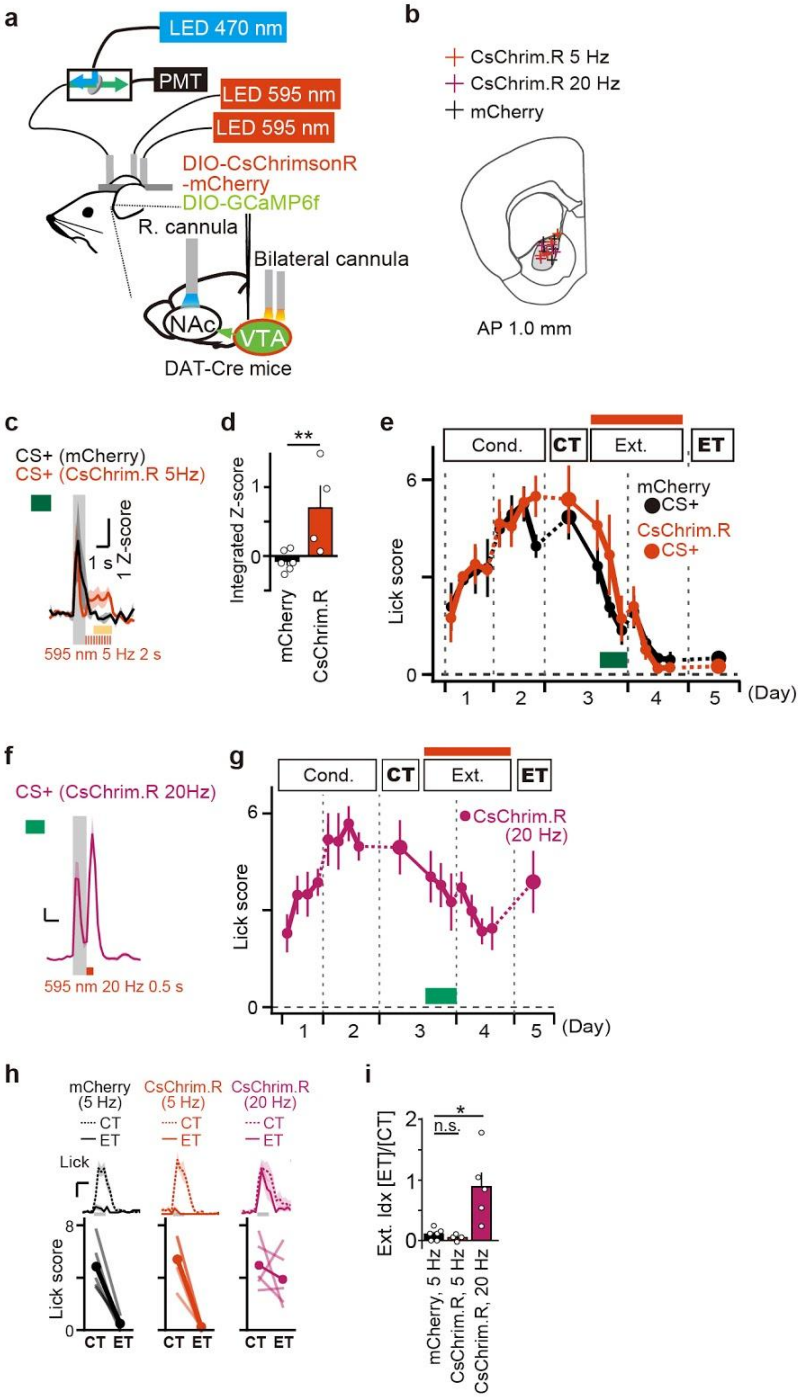

Figure 12. Optogenetic DA stimulation during extinction learning.

a, Schematic of experimental setup for optogenetic stimulation of DA neurons under monitoring with fiber photometry during extinction. AAVs were injected to express CsChrimsonR and GCaMP6f in VTA-DA neurons. Optical fibres were implanted at the VTA for stimulation of CsChrimsonR and at the right NAc for photometry of GCaMP. **b**, Locations of the fibre tips plotted on the reference atlas. **c**, Averaged *z*-scores of

photometry responses on 5 Hz stimulation in the extinction period indicated with the green bar in **e**. **d,** Integrated *z*-score of photometry responses 1.5–3 s after CS onset (**c**, orange bar). ***P* = 0.004 by Two-sided Mann–Whitney U test. **e**, Time course of lick scores. Red bar indicates the period of 5 Hz optogenetic stimulation. CT, conditioning test; ET, extinction test. **f,** Averaged *z*-scores of photometry responses on 20 Hz stimulation in the extinction period indicated with the green bar in **g. g**, Time course of lick scores. Red bar indicates the period of 20 Hz optogenetic stimulation. **h**, Averaged lick traces (top) and lick scores (bottom) at conditioning test and extinction test. **i**, Extinction indices (Ext. index: (lick score at ET) / (lick score at CT)). *n* = 6 (mCherry, 5 Hz), 4 (CsChrimsonR, 5 Hz) and 5 mice (CsChrimsonR, 20 Hz). Kruskal–Wallis test (*P* = 0.009; post hoc Steel's test, CsChrimsonR 5 Hz versus mCherry, *P* = 0.34; CsChrimsonR 20 Hz versus mCherry, **P* = 0.02). The numbers of mice are biological replicates.

LTP-related signals of D2-SPN in discrimination learning

Given that DA dips in the nucleus accumbens are detected by D2Rs to disinhibit LTP in D2-SPNs, which is dependent on CaMKII[3], I addressed whether this signal in D2-SPNs were also required for discrimination learning. When autocamtide-2 inhibitory peptide (AIP)[39], a CaMKII inhibitory peptide, was expressed in D2-SPNs at the lateral NAc core using a viral vector (Fig. 13a, b), discrimination learning was blocked (Fig. 13c–f and 14). AIP expression in D2-SPNs did not affect generalization (Fig. 13d), supporting the contention that generalized conditioning is selectively mediated by D1-SPNs. Of note, AIP expression did not significantly affect extinction learning (Fig. 15), consistent with the results that 5-Hz DA stimulation did not inhibit it (Fig. 12a-e).

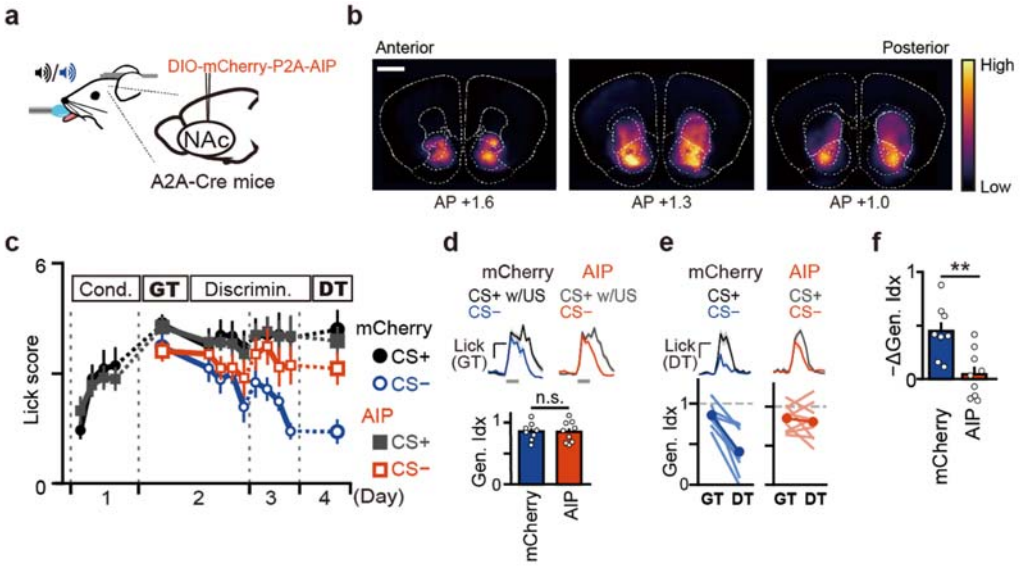

Figure 13. Effect of AIP expression in D2-SPNs on discrimination learning.
a, Schematic of virus injection. **b**, Averaged intensity of AIP expression for the mice used in the experiment (n = 9 mice). Reference atlases are overlaid. **c**, Time course of lick scores. GT, generalization test; DT, discrimination test. **d**, Averaged lick traces (top) and generalization indices (bottom) at generalization test. **e**, Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **f**, ΔGeneralization indices. $^{**}P = 0.006$ by Two-sided Mann–Whitney U test. $n = 8$ (mCherry) and 9 (AIP) mice. The numbers of mice are biological replicates.

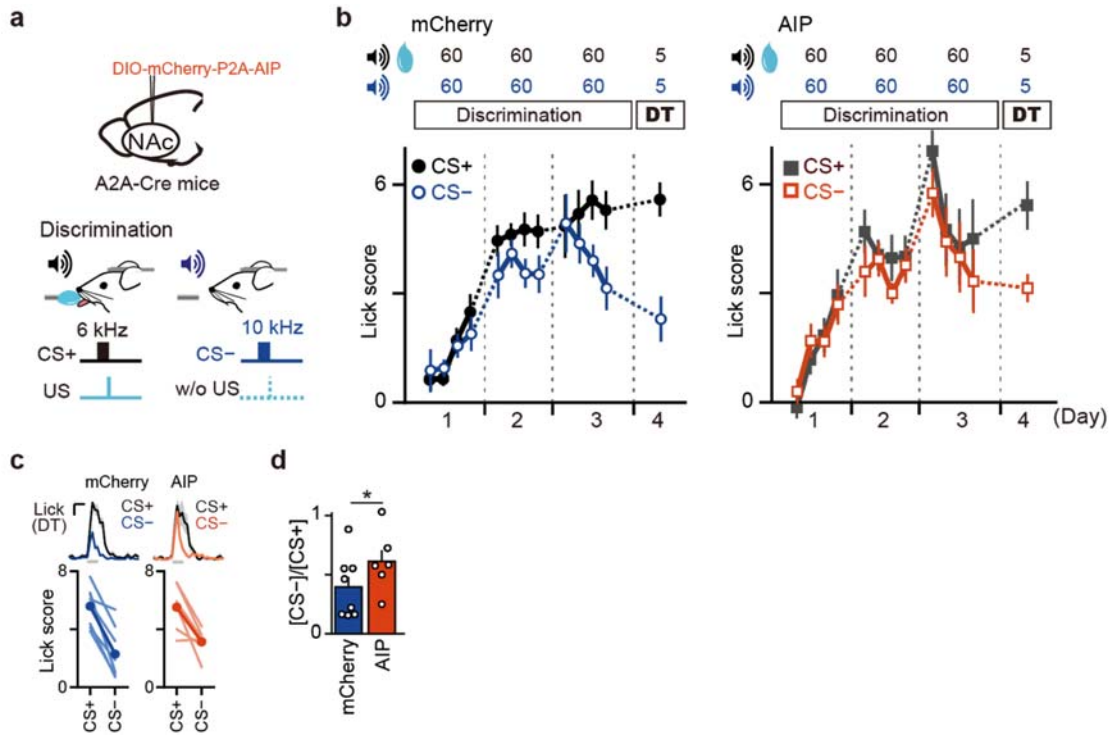

Figure 14. Effect of AIP expression in D2-SPNs on conventional discrimination learning.

a, Schematic of virus injection and conventional discrimination paradigm (Figure 2). **b**, Numbers of trials during the discrimination task (top) and time course of lick scores (bottom). DT, discrimination test. **c**, Averaged lick traces (top) and lick scores (bottom) at discrimination test. **f**, (lick score in CS− trials) / (lick score in CS+ trials) at discrimination test. **P* = 0.041 by Two-sided Mann–Whitney U test. *n* = 8 (mCherry) and 6 (AIP) mice. The numbers of mice are biological replicates.

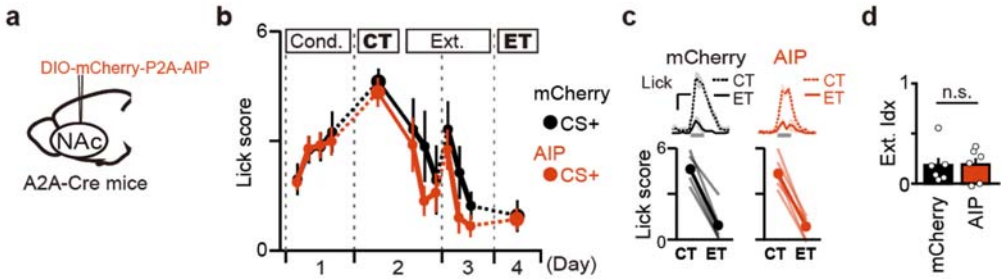

Figure 15. Effect of AIP expression in D2-SPNs on extinction learning.
a, Schematic of virus injection. **b**, Time course of lick scores. CT, conditioning test; ET, extinction test. **c**, Averaged lick traces (top) and lick scores (bottom) at conditioning test and extinction test. **d**, Extinction indices (Ext. index: (lick score at ET) / (lick score at CT)). $P = 0.82$ by wo-sided Mann–Whitney U test. $n = 6$ (mCherry) and 6 (AIP) mice. The numbers of mice are biological replicates.

As chronic expression of AIP may induce some collateral effects, I tested whether the learning was also modulated by acute manipulation of A2AR and D2R sinallings, which was required for and inhibited LTP in D2-SPNs, respectively[3]. Infusion of the A2AR antagonist SCH58261 into the bilateral NAc (Fig. 16a) during the discrimination period blocked discrimination learning (measured on day 4 without the drug; Fig. 16b–d), suggesting that signals were required during learning. By contrast, infusion of the D2R antagonist sulpiride, which facilitated LTP in the previous ex vivo study[3], facilitated one-day discrimination learning (Fig. 17). Evaluation of AIP expression (Fig. 13b) and estimated spread of drugs (Fig. 18) suggested that the NAc core had a major role in these effects.

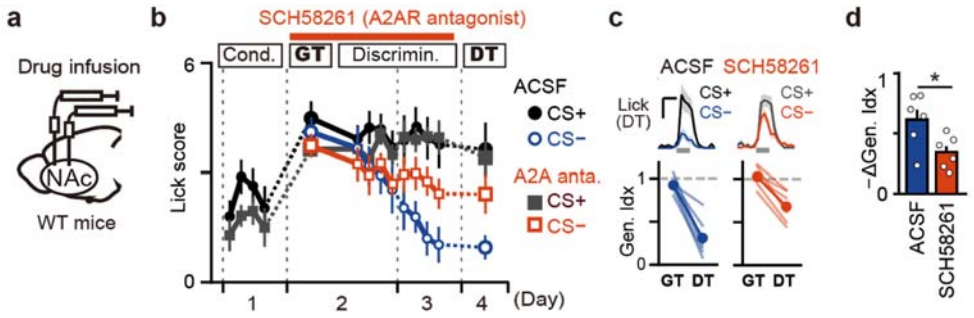

Figure 16. Effect of A2AR antagonist on discrimination learning.
a, Schematic of drug infusion. **b**, Time course of lick scores. Red bar indicates the period of drug infusion. GT, generalization test; DT, discrimination test. **c**, Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **f**, ΔGeneralization indices. **P* = 0.041 by Two-sided Mann–Whitney U test. *n* = 6 (ACSF) and 6 (SCH58261) mice. The numbers of mice are biological replicates.

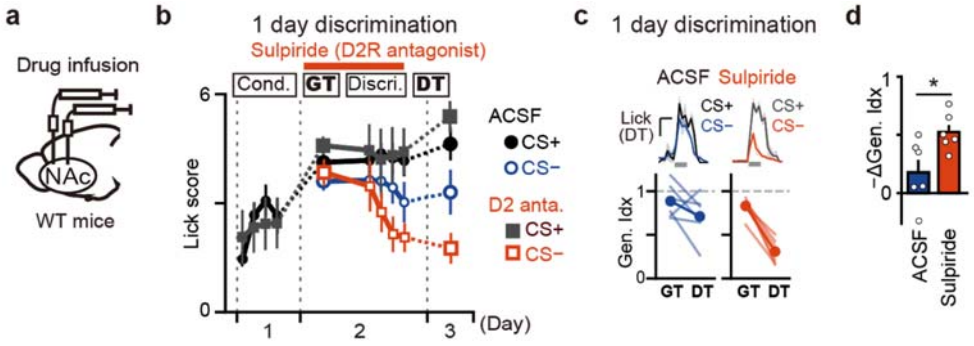

Figure 17. Effect of D2R antagonist on discrimination learning.
a, Schematic of drug infusion. **b**, Time course of lick scores. Red bar indicates the period of drug infusion. GT, generalization test; DT, discrimination test. **c**, Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **f**, ΔGeneralization indices. **P* = 0.041 by Two-sided Mann–Whitney U test. *n* = 6 (ACSF) and 6 (Sulpiride) mice. The numbers of mice are biological replicates.

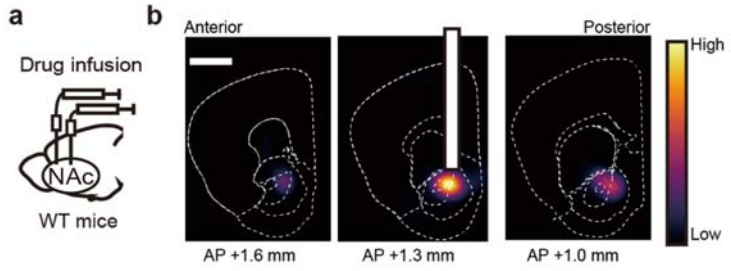

Figure 18. Estimation of drug spread.

a, Schematic of drug infusion. **b**, Averaged intensity of a fluorescent dye (Alexa 594, n = 3 mice) that was injected through the unilateral cannula over the same period as in the discrimination tasks with overlaid reference atlases. Scale bar, 500 μm. The numbers of mice are biological replicates.

Methamphetamine effects

Finally, I investigated whether the discrimination learning mediated by D2Rs was impaired by repeated exposure to methamphetamine (MAP), which causes psychosis in humans and sensitization in mice (Fig. 19)[40,41]. When MAP (1 mg kg^{-1}) was injected intraperitoneally 30 min before the discrimination task, discrimination was impaired as shown in the discrimination test (Fig. 20). I could not engage agitated mice in the task immediately after the injection of 1 mg kg^{-1} MAP, nor 30 min after the injection of 4 mg kg^{-1} of MAP. Then, the effects of MAP on extracellular DA dynamics were examined by fibre photometry of the DA sensor $GRAB_{DA1m}$[30] expressed in the lateral NAc core (Fig. 21). Consistent with the inhibition of dopamine transporter (DAT) by MAP, administration of MAP induced a gradual, dose-dependent increase in $GRAB_{DA1m}$ signals in the NAc over 30 min, followed by a gradual decrease (Fig. 22). During the discrimination task (Fig. 23a-c), DA dips were reflected in $GRAB_{DA1m}$ signals in the NAc in control, saline-treated mice (Fig. 23d), similarly to GCaMP6f signals, indicating that DA neurons' axonal activity and DA concentration were coupled during this task. MAP administration reduced the size of the DA dips (Fig. 23d), while maintaining DA responses to the US (Fig. 23e) during the discrimination task, and

prevented the suppression of the licking responses (Fig. 23a-c) and the DA responses to the CS− measured at the discrimination test (Fig. 23f, g).

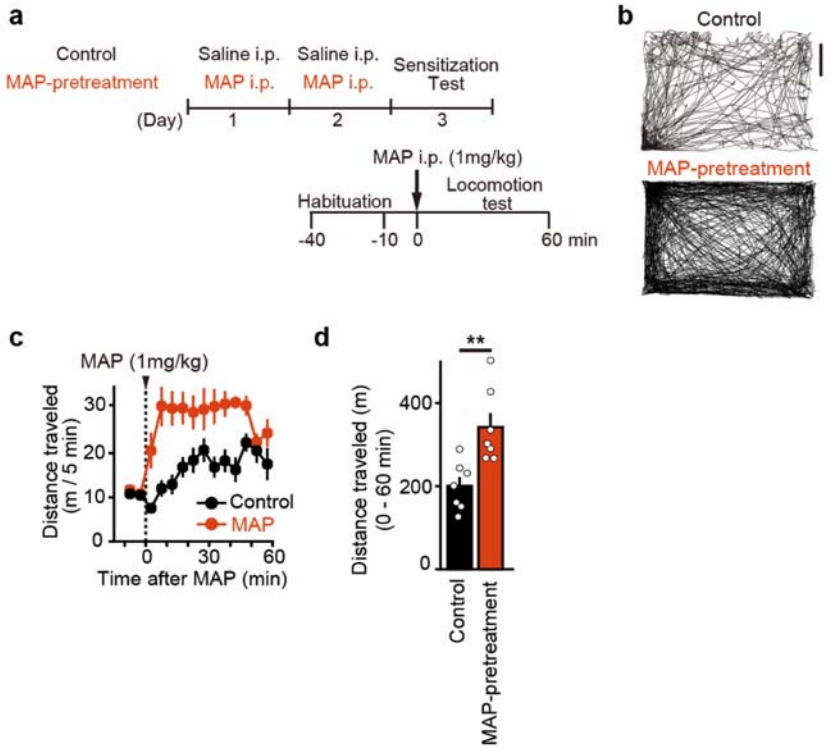

Figure 19. Effect of MAP on locomotor activity.
a, Experimental schedule for locomotor sensitization by MAP (1 mg kg^{-1}). **b**, Representative locomotion traces 0–15 min after MAP administration in sensitization test. Scale bar, 10 cm. **c**, Time course of average distances travelled in sensitization tests against time. $n = 7$ (control) and 7 mice (MAP). **d**, Plots of average travel distances by control mice ($n = 7$) and MAP-treated mice ($n = 7$). $^{**}P = 0.004$ by two-sided Mann–Whitney U test. The numbers of mice are biological replicates.

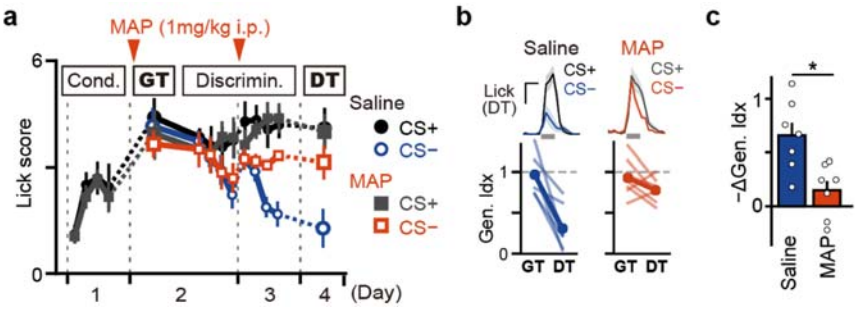

Figure 20. Effect of MAP on discrimination learning.

a, Time course of lick scores. MAP (1 mg kg^{-1}, intraperitoneal) was administered 30 min before the discrimination task on days 2 and 3 (red arrowheads). GT, generalization test; DT, discrimination test. **b**, Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **c**, ΔGeneralization indices. $^{**}P = 0.0099$ by two-sided Mann–Whitney *U* test. $n = 7$ mice per group. The numbers of mice are biological replicates.

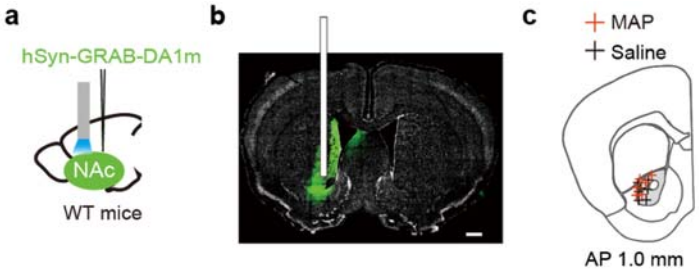

Figure 21. Methods for measurement of extracellular DA dynamics.
a, Schematic of virus injection. **b**, **c**, Representative confocal image of $GRAB_{DA1m}$ expression (**b**) and locations of the fibre tips for the mice used in the experiment (Fig. 22) (**c**). Grey area indicates NAc core. Scale bar, 500 μm.

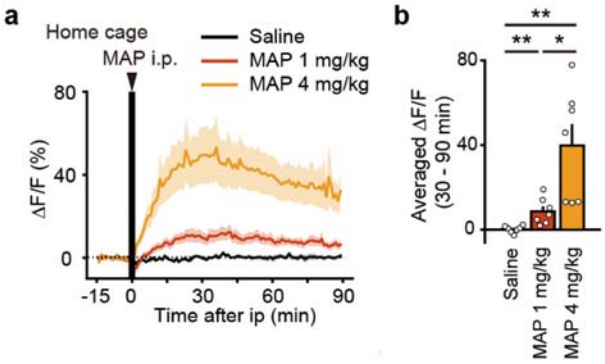

Figure 22. Effect of MAP on dopamine concentration.
a, Time course of $GRAB_{DA1m}$ responses to MAP administration. Saline, 1 mg kg^{-1} MAP, or 4 mg kg^{-1} MAP were administered intraperitoneally in the home cage. $n = 7$ mice. Each experiment was conducted on a separate day and the order of drug administration was counterbalanced. **b**, Plots of $GRAB_{DA1m}$ responses 30–90 min after MAP injection. $n = 7$ mice. Friedman's test ($P = 9.0 \times 10^{-4.}$). Post hoc Steel–Dwass test (saline versus MAP 1 mg kg^{-1}, $^{**}P = 0.008$; saline versus MAP 4 mg kg^{-1}, $^{**}P = 0.005$; MAP 1 mg kg^{-1} versus MAP 4 mg kg^{-1}, $^{*}P = 0.048$). The numbers of mice are biological replicates.

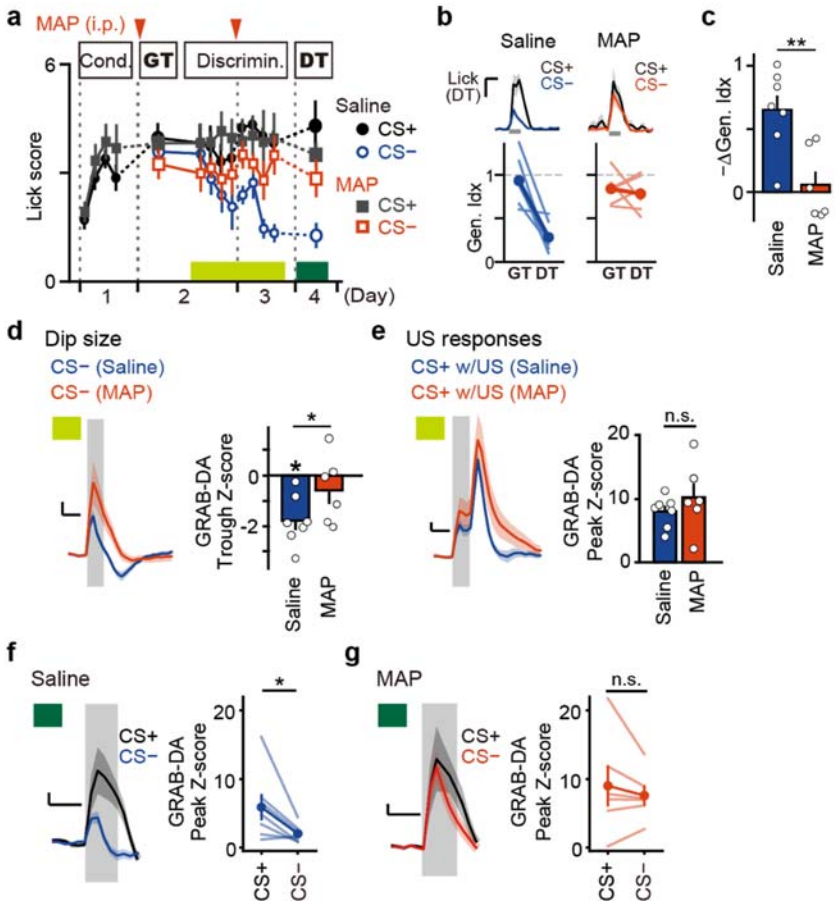

Figure 23. Effect of MAP on dopamine concentration during discrimination learning.

a, Time course of lick scores for mice injected with $GRAB_{DA1m}$ and treated with MAP. MAP was administered in the same way as in Fig. 19a. **b**, Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **c**, ΔGeneralization indices. $^{**}P = 0.005$ by Two-sided Mann–Whitney *U* test. **d**, **e**, $GRAB_{DA1m}$ responses to CS− (**d**) and CS+ with US (**e**) averaged over discrimination task on days 2 and 3. DA dips in CS− trials during discrimination task (**d**) were tested by two-sided Wilcoxon signed-rank test versus 0 (saline, $^{*}P = 0.016$; MAP, $P = 0.63$). Dip size was evaluated by the *z*-score trough during 2–3 s after CS onset to avoid contamination by decay of CS responses. Effects of MAP on dip size (**d**) and phasic DA responses to US (**e**) were tested by two-sided Mann–Whitney *U* test (**d**, $^{*}P = 0.035$; **e**, $P = 0.25$). Data were sampled during the discrimination sessions as indicated with the light green bar in **a**. **f**, **g**, $GRAB_{DA1m}$ responses to CS+ and CS− at discrimination test from mice injected with saline (**f**) or MAP (**g**). Two-sided Wilcoxon signed-rank test (saline, $n = 7$ mice, $^{*}P = 0.031$; MAP, $n = 6$ mice, $P = 0.56$). Data were sampled during the discrimination test as indicated with the green bar in **a**. $n = 7$ (saline) and 6 (MAP) mice. The numbers of mice are biological replicates.

Of note, conditioned place preference (CPP)[42], an addiction-related behaviour involving D1R mechanisms[43], was not formed by the MAP treatment under the conditions used here (1 mg kg^{-1}, 30 min before the task; Fig. 24), although a previous study showed that conditioning immediately after 1 mg kg^{-1} MAP administration established CPP[42]. These data suggest that discrimination learning can be more sensitive to MAP treatment than D1Rs-related behaviours.

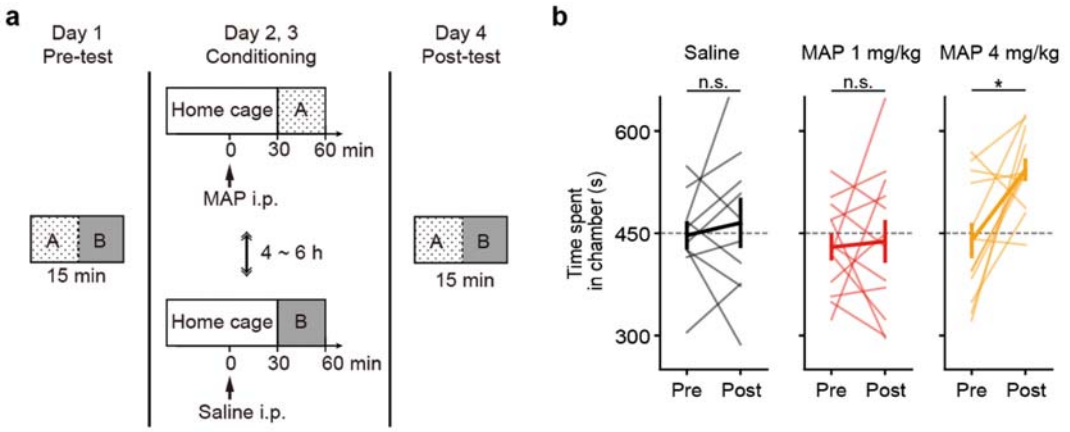

Figure 24. Conditioned place preference test with delays after MAP administration.

a, Schematic of CPP protocol with delays after MAP administration. Allocation of chambers A and B and the order of MAP and saline injection were counterbalanced. **b**, Plots of time spent in chambers. Two-sided Wilcoxon signed-rank test (saline, $n = 10$ mice, $P = 0.77$; MAP 1 mg kg^{-1}, $n = 12$ mice, $P = 0.91$; MAP 4 mg kg^{-1}, $n = 12$ mice, $^{*}P = 0.012$). The numbers of mice are biological replicates.

Infusion of a D2R antagonist into the NAc of MAP-treated mice during the discrimination period facilitated the discrimination learning (Fig. 25a, c, d). By contrast, administration of a D2R antagonist at the discrimination test did not affect the extent of discrimination (Fig. 25b-d), suggesting that the D2R antagonist facilitated discrimination learning but not its expression.

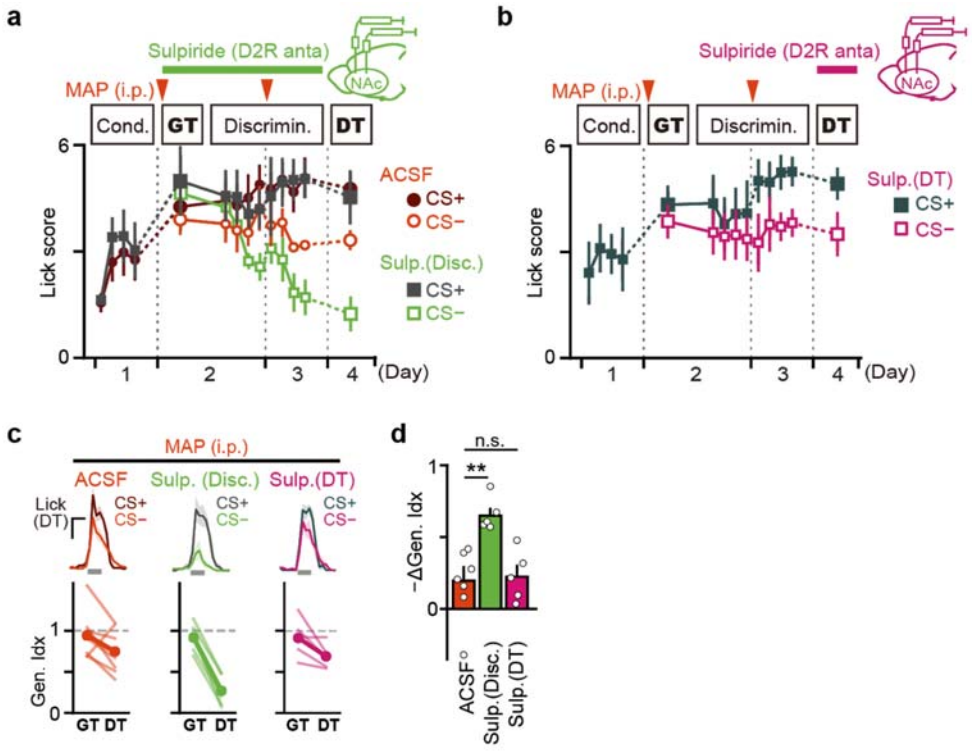

Figure 25. Effect of D2R antagonist on discrimination learning of MAP-treated mice.

a, b, Time course of lick scores of mice treated with MAP and infused with sulpiride during discrimination learning (**a**) or at discrimination test (**b**). MAP was administered as in Fig. 19a. Green (**a**) and pink bar (**b**) indicate the period of bilateral infusion of sulpiride into the NAc. $n = 7$ (ACSF), 5 (sulpiride during discrimination) and 5 mice (sulpiride at discrimination test). Through days 2 to 4, either ACSF or sulpiride was infused for all the conditions to control the manipulation of infusion. **c**, Averaged lick traces at discrimination test (top) and generalization indices at generalization test and discrimination test (bottom). **d**, ΔGeneralization indices. Kruskal–Wallis test ($P = 0.007$; post hoc Steel's test, ACSF versus sulpiride during discrimination, $^{**}P = 0.009$; ACSF versus sulpiride in discrimination test, $P = 0.99$). The numbers of mice are biological replicates.

Discussion

Reward learning involves marked generalization of stimuli[10,13,36], which are probably beneficial to effectively predict rewards in the real world, where reward-predicting stimuli are likely to vary. In this study, I demonstrated the generalized conditioning was via D1Rs in the NAc (Fig. 3), consistent with the broad auditory tuning of striatal neurons[44]. When a generalized cue was actually not followed by a reward, DA dips occurred at the lateral part of the NAc core (Fig. 4, 5). Bidirectional optogenetic manipulation under the monitoring with fiber photometry showed the causal link between these DA dips and the discrimination learning (Fig. 9, 10). This alone doesn't exclude the possibility that the mice learned discrimination using dopamine signals in other brain areas than the NAc, because the optogenetic stimulation targets the somata of dopamine neurons which projects to various areas. However, the experiments using inhibitory peptide expression (Fig. 13, 14) and drug infusion (Fig. 16, 17) locally in the NAc showed that discrimination learning was dependent on LTP-related signals of D2-SPNs there. These results, taken together with the prior ex vivo study showing that DA dips are detected by D2Rs to disinhibit LTP in D2-SPNs[3], suggest that DA dips in the lateral NAc control discrimination learning via D2Rs.

The multi-site photometry recordings revealed distinct DA neuronal activity patterns at the lateral and medial NAc, the latter of which showed elongated burst responses to reward and no significant DA dips upon reward omission (Fig. 5c). Such heterogeneity of DA neurons axonal activities even within the NAc, in addition to the well described one across dorsal-ventral axis of the striatum[17,45], suggest elaborate regional tuning to control the specific functions of each axonal target region. Technical advances allowing selective manipulation of DA axonal activities in each region under monitoring should help to determine such specific functions.

Contrary to theoretical predictions[46–48], extinction required neither DA dips nor LTP-related signals of D2-SPNs in the NAc (Fig. 12a-e, 15). Although a prior study showed optogenetic stimulation of dopamine neurons slowed extinction learning[23], the stimulation in this study was at 20 Hz, as my experiments in Fig. 12f-g, which would induce DA bursts and involve D1Rs signals. Also, a prior genetic lesion study showed that extinction of operant conditioning was dependent on D2-SPNs in the striatum[49]. Although my results do not exclude the possibility that D2-SPNs' activities are necessary for extinction, they do suggest that extinction does not involve DA dips or LTP in D2-SPNs in the NAc. This opens up new questions as to how different forms of

negative prediction errors relate to dopamine signals and its downstream circuits. It is possible that general reward omission, contrary to reward omission in rewarding contexts, involves mechanisms such as switching of a contextual representation[50], whereby DA firing no longer encodes the prediction error. It would be of interest to explore the neural substrates to enable such operation.

Discrimination learning was vulnerable to dysregulation of DA signalling, as exemplified by MAP treatment, which led to inhibiton of DA dips during discrimination learning and the failure to suppress behavioral (Fig. 20b, c and 23b, c) nor neuronal (Fig. 23f, g) responses to the CS−. This is consistent with the collaborator's ex vivo study showing that MAP treatment impaired DA dips and subsequent induction of LTP in D2-SPNs[3]. The D2R antagonist facilitated the discrimination learning (Fig. 25a, c,d), as well as LTP in D2-SPNs ex vivo[3]. These results suggest that the impairments in discrimination learning via D2R may underlie some of the neuropsychiatric symptoms associated with dysregulation of D2Rs, such as the psychotic symptoms seen in stimulants users[40,51] and schizophrenic patients[52], particularly when they have to discriminate in a novel context after life events, a known trigger for psychosis [53]. In fact, a functional MRI study of individuals with schizophrenia showed augmented NAc

activity in response to the CS− after conditioning[48], supporting the aberrant salience hypothesis, whereby over-attribution of salience to irrelevant environmental events underlies delusional symptoms[54]. Investigation of the impairments in discrimination learning and its progression in people with schizophrenia might provide novel early diagnostic methods.

Also, the effect of the D2R antagonists, or antipsychotics, was positive only when they were treated during discrimination task, but not when treated at discrimination test (Fig. 25b), suggesting that D2R antagonists facilitate discrimination learning but not its expression. This leads to a model that these drugs gradually ameliorate symptoms by assisting discrimination learning through real-world experiences over days, which is consistent with the therapeutic delay of D2Rs-targeting antipsychotics in clinical practice[52] . Considering the delusional symptoms in schizophrenia are often persecutory, examining the role of DA dips and D2Rs and the effect of antipychotics in discrimination of aversive stimuli would advance our understanding of the biological underpinnings of schizophrenia and related behaviors.

The current results suggest that DA dips, via the detection system by D2Rs, refine the generalized reward learning mediated by D1Rs. Such cooperative tuning of conditioning by D1Rs and D2Rs should play fundamental roles in appropriate adaptation to the environment.

Acknowledgements

I would like to express my sincere gratitude to Prof. Haruo Kasai and Dr. Sho Yagishita for their invaluable mentoring, consistent support and patience. Further, I would like to thank all the collaborators, especially to Dr. Yusuke Iino and Dr. Kenji Yamaguchi, for considerate guidance. I am also grateful to Ms. Arisa Kurabayashi, Mr. Masayasu Asaumi and Ms. Atena Nishikawa for their technical assistance and all the members of the Structural Physiology lab for their understanding and help. To conclude, I cannot forget to thank my family and friends for all the support during this research.

References

1. Creese, I., Burt, D. R. & Snyder, S. H. Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science* **192,** 481–483 (1976).
2. Seeman, P. Targeting the dopamine D2 receptor in schizophrenia. *Expert Opin. Ther. Targets* **10,** 515–531 (2006).
3. Iino, Y., Sawada, T., Yamaguchi, K., Tajiri, M., Ishii, S., Kasai, H. & Yagishita, S. Dopamine D2 receptors in discrimination learning and spine enlargement. *Nature* **579,** 555–560 (2020).
4. Tan, K. R., Yvon, C., Turiault, M., Mirzabekov, J. J., Doehner, J., Labouèbe, G., Deisseroth, K., Tye, K. M. & Lüscher, C. GABA neurons of the VTA drive conditioned place aversion. *Neuron* **73,** 1173–1183 (2012).
5. Danjo, T., Yoshimi, K., Funabiki, K., Yawata, S. & Nakanishi, S. Aversive behavior induced by optogenetic inactivation of ventral tegmental area dopamine neurons is mediated by dopamine D2 receptors in the nucleus accumbens. *Proc Natl Acad Sci USA* **111,** 6455–6460 (2014).
6. Chang, C. Y., Esber, G. R., Marrero-Garcia, Y., Yau, H.-J., Bonci, A. & Schoenbaum, G. Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. *Nat. Neurosci.* **19,** 111–116 (2016).
7. Sutton, R. S. Learning to predict by the methods of temporal differences. *Mach. Learn.* **3,** 9–44 (1988).
8. Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482,** 85–88 (2012).

9. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* **72,** 1024–1027 (1994).
10. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412,** 43–48 (2001).
11. Pan, W.-X., Schmidt, R., Wickens, J. R. & Hyland, B. I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* **25,** 6235–6242 (2005).
12. D'Ardenne, K., McClure, S. M., Nystrom, L. E. & Cohen, J. D. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* **319,** 1264–1267 (2008).
13. Day, J. J., Roitman, M. F., Wightman, R. M. & Carelli, R. M. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* **10,** 1020–1028 (2007).
14. Hart, A. S., Rutledge, R. B., Glimcher, P. W. & Phillips, P. E. M. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *J. Neurosci.* **34,** 698–704 (2014).
15. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47,** 129–141 (2005).
16. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275,** 1593–1599 (1997).
17. Menegas, W., Babayan, B. M., Uchida, N. & Watabe-Uchida, M. Opposite initialization to novel cues in dopamine signaling in ventral and posterior striatum in mice. *elife* **6,** (2017).
18. Björklund, A. & Dunnett, S. B. Dopamine neuron systems in the brain: an update.

Trends Neurosci. **30,** 194–202 (2007).

19. Patriarchi, T., Cho, J. R., Merten, K., Howe, M. W., Marley, A., Xiong, W.-H., Folk, R. W., Broussard, G. J., Liang, R., Jang, M. J., Zhong, H., Dombeck, D., von Zastrow, M., Nimmerjahn, A., Gradinaru, V., Williams, J. T. & Tian, L. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* **360,** (2018).
20. Hamid, A. A., Pettibone, J. R., Mabrouk, O. S., Hetrick, V. L., Schmidt, R., Vander Weele, C. M., Kennedy, R. T., Aragona, B. J. & Berke, J. D. Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19,** 117–126 (2016).
21. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321,** 848–851 (2008).
22. Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G. C. R., Urakubo, H., Ishii, S. & Kasai, H. A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* **345,** 1616–1620 (2014).
23. Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K. & Janak, P. H. A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16,** 966–973 (2013).
24. Tsai, H.-C., Zhang, F., Adamantidis, A., Stuber, G. D., Bonci, A., de Lecea, L. & Deisseroth, K. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324,** 1080–1084 (2009).
25. Yamaguchi, K., Maeda, Y., Sawada, T., Iino, Y., Tajiri, M., Nakazato, R., Kasai, H. & Yagishita, S. The minimal behavioral time window for reward conditioning in the nucleus accumbens of mice. *BioRxiv* (2019). doi:10.1101/641365
26. Saunders, B. T., Richard, J. M., Margolis, E. B. & Janak, P. H. Dopamine neurons

create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nat. Neurosci.* **21,** 1072–1083 (2018).

27. Zweifel, L. S., Parker, J. G., Lobb, C. J., Rainwater, A., Wall, V. Z., Fadok, J. P., Darvas, M., Kim, M. J., Mizumori, S. J. Y., Paladini, C. A., Phillips, P. E. M. & Palmiter, R. D. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proc Natl Acad Sci USA* **106,** 7281–7288 (2009).
28. Klapoetke, N. C., Murata, Y., Kim, S. S., Pulver, S. R., Birdsey-Benson, A., Cho, Y. K., Morimoto, T. K., Chuong, A. S., Carpenter, E. J., Tian, Z., Wang, J., Xie, Y., Yan, Z., Zhang, Y., Chow, B. Y., Surek, B., Melkonian, M., Jayaraman, V., Constantine-Paton, M., Wong, G. K.-S. & Boyden, E. S. Independent optical excitation of distinct neural populations. *Nat. Methods* **11,** 338–346 (2014).
29. Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nat. Protoc.* **1,** 1412–1428 (2006).
30. Sun, F., Zeng, J., Jing, M., Zhou, J., Feng, J., Owen, S. F., Luo, Y., Li, F., Wang, H., Yamaguchi, T., Yong, Z., Gao, Y., Peng, W., Wang, L., Zhang, S., Du, J., Lin, D., Xu, M., Kreitzer, A. C., Cui, G. & Li, Y. A genetically encoded fluorescent sensor enables rapid and specific detection of dopamine in flies, fish, and mice. *Cell* **174,** 481-496.e19 (2018).
31. Guo, Q., Zhou, J., Feng, Q., Lin, R., Gong, H., Luo, Q., Zeng, S., Luo, M. & Fu, L. Multi-channel fiber photometry for population neuronal activity recording. *Biomed. Opt. Express* **6,** 3919–3931 (2015).
32. Franklin, K. B. J. & Paxinos, G. *The mouse brain in stereotaxic coordinates*. (2013).

33. Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W. & Bethge, M. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21,** 1281–1289 (2018).

34. de Jong, J. W., Afjei, S. A., Pollak Dorocic, I., Peck, J. R., Liu, C., Kim, C. K., Tian, L., Deisseroth, K. & Lammel, S. A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. *Neuron* **101,** 133-151.e7 (2019).

35. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459,** 837–841 (2009).

36. Beyeler, A., Namburi, P., Glober, G. F., Simonnet, C., Calhoon, G. G., Conyers, G. F., Luck, R., Wildes, C. P. & Tye, K. M. Divergent Routing of Positive and Negative Information from the Amygdala during Memory Retrieval. *Neuron* **90,** 348–361 (2016).

37. Marcott, P. F., Mamaligas, A. A. & Ford, C. P. Phasic dopamine release drives rapid activation of striatal D2-receptors. *Neuron* **84,** 164–176 (2014).

38. Eshel, N., Bukwich, M., Rao, V., Hemmelder, V., Tian, J. & Uchida, N. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* **525,** 243–246 (2015).

39. Murakoshi, H., Shin, M. E., Parra-Bueno, P., Szatmari, E. M., Shibata, A. C. E. & Yasuda, R. Kinetics of endogenous camkii required for synaptic plasticity revealed by optogenetic kinase inhibitor. *Neuron* **94,** 37-47.e5 (2017).

40. Ujike, H. & Sato, M. Clinical features of sensitization to methamphetamine observed in patients with methamphetamine dependence and psychosis. *Ann. N. Y. Acad. Sci.* **1025,** 279–287 (2004).

41. Robinson, T. E. & Becker, J. B. Enduring changes in brain and behavior produced by chronic amphetamine administration: a review and evaluation of animal models of amphetamine psychosis. *Brain Res.* **396,** 157–198 (1986).

42. K Shimosato, S. O. Simultaneous monitoring of conditioned place preference and locomotor sensitization following repeated administration of cocaine and methamphetamine. *Pharmacol. Biochem. Behav.* **66,** 285–292 (2000).

43. Lobo, M. K., Covington, H. E., Chaudhury, D., Friedman, A. K., Sun, H., Damez-Werno, D., Dietz, D. M., Zaman, S., Koo, J. W., Kennedy, P. J., Mouzon, E., Mogri, M., Neve, R. L., Deisseroth, K., Han, M.-H. & Nestler, E. J. Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* **330,** 385–390 (2010).

44. Bordi, F. & LeDoux, J. Sensory tuning beyond the sensory system: an initial analysis of auditory response properties of neurons in the lateral amygdaloid nucleus and overlying areas of the striatum. *J. Neurosci.* **12,** 2493–2503 (1992).

45. Howe, M. W. & Dombeck, D. A. Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature* **535,** 505–510 (2016).

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

47. Samson, R. D., Frank, M. J. & Fellous, J.-M. Computational models of reinforcement learning: the role of dopamine as a reward signal. *Cogn. Neurodyn.* **4,** 91–105 (2010).

48. Redish, A. D., Jensen, S., Johnson, A. & Kurth-Nelson, Z. Reconciling

reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychol. Rev.* **114,** 784–805 (2007).

49. Matamales, M., McGovern, A. E., Mi, J. D., Mazzone, S. B., Balleine, B. W. & Bertran-Gonzalez, J. Local D2- to D1-neuron transmodulation updates goal-directed learning in the striatum. *Science* **367,** 549–555 (2020).
50. Gershman, S. J., Jones, C. E., Norman, K. A., Monfils, M.-H. & Niv, Y. Gradual extinction prevents the return of fear: implications for the discovery of state. *Front. Behav. Neurosci.* **7,** 164 (2013).
51. Farnia, V., Shakeri, J., Tatari, F., Juibari, T. A., Yazdchi, K., Bajoghli, H., Brand, S., Abdoli, N. & Aghaei, A. Randomized controlled trial of aripiprazole versus risperidone for the treatment of amphetamine-induced psychosis. *Am. J. Drug Alcohol Abuse* **40,** 10–15 (2014).
52. Agid, O., Seeman, P. & Kapur, S. The “delayed onset” of antipsychotic action--an idea whose time has come and gone. *J. Psychiatry Neurosci.* **31,** 93–100 (2006).
53. Beards, S., Gayer-Anderson, C., Borges, S., Dewey, M. E., Fisher, H. L. & Morgan, C. Life events and psychosis: a review and meta-analysis. *Schizophr. Bull.* **39,** 740–747 (2013).
54. Howes, O. D. & Nour, M. M. Dopamine and the aberrant salience hypothesis of schizophrenia. *World Psychiatry* **15,** 3–4 (2016).