Title

From Prediction to Action: Dissociable Roles of Ventral Tegmental Area and Substantia Nigra Dopamine Neurons in Instrumental Reinforcement.

Journal

Journal of Neuroscience, 43(21)

Authors

Pribut, Heather

Janak, Patricia

Keiflin, Ronald

et al.

Publication Date

2023-05-24

DOI

10.1523/JNEUROSCI.0028-23.2023

Peer reviewed

Behavioral/Cognitive

From Prediction to Action: Dissociable Roles of Ventral Tegmental Area and Substantia Nigra Dopamine Neurons in Instrumental Reinforcement

Kurt M. Fraser,[1] Heather J. Pribut,[1] Patricia H. Janak,[1,2,3] and Ronald Keiflin[4,5]

[1]Department of Psychological & Brain Sciences, Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, Maryland 21218, [2]The Solomon H. Snyder Department of Neuroscience, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, Maryland 21205, [3]Kavli Neuroscience Discovery Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, [4]Department of Psychological & Brain Sciences, University of California, Santa Barbara, Santa Barbara, California 93106, and [5]Neuroscience Research Institute, University of California Santa Barbara, Santa Barbara, California, 93106

Reward seeking requires the coordination of motor programs to achieve goals. Midbrain dopamine neurons are critical for reinforcement, and their activation is sufficient for learning about cues, actions, and outcomes. Here we examine in detail the mechanisms underlying the ability of ventral tegmental area (VTA) and substantia nigra (SNc) dopamine neurons to support instrumental learning. By exploiting numerous behavioral tasks in combination with time-limited optogenetic manipulations in male and female rats, we reveal that VTA and SNc dopamine neurons generate reinforcement through separable psychological processes. VTA dopamine neurons imbue actions and their associated cues with motivational value that allows flexible and persistent pursuit, whereas SNc dopamine neurons support time-limited, precise, action-specific learning that is nonscalable and inflexible. This architecture is reminiscent of actor–critic reinforcement learning models with VTA and SNc instructing the critic and actor, respectively. Our findings indicate that heterogeneous dopamine systems support unique forms of instrumental learning that ultimately result in disparate reward-seeking strategies.

Key words: **dopamine; ICSS; motivation; operant conditioning; reinforcement; reward**

Significance Statement

Dopamine neurons in the midbrain are essential for learning, motivation, and movement. Here we describe in detail the ability of VTA and SNc dopamine neurons to generate instrumental reinforcement, a process where an agent learns about actions they can emit to earn reward. While rats will avidly work and learn to respond for activation of VTA and SNc dopamine neurons, we find that only VTA dopamine neurons imbue actions and their associated cues with motivational value that spur continued pursuit of reward. Our data support a hypothesis that VTA and SNc dopamine neurons engage distinct psychological processes that have consequences for our understanding of these neurons in health and disease.

Introduction

Midbrain dopamine (DA) neurons encode reward prediction errors, a fundamental parameter in reinforcement learning (Schultz et al., 1997; Waelti et al., 2001; Glimcher, 2011), and their activation promotes learning about events leading to reward (Steinberg et al., 2013; Ilango et al., 2014; Ramayya et al., 2014; Chang et al., 2016; Stauffer et al., 2016; Keiflin et al., 2019). Although early studies reported relatively broad and homogeneous responses to unexpected rewards across midbrain dopamine cell groups—including the ventral tegmental area (VTA) and the substantia nigra (SNc)—(Mirenowicz and Schultz, 1996; Schultz, 1998; Bayer and Glimcher, 2005), more recent studies have demonstrated considerable heterogeneity in the response pattern of different dopamine subsystems, particularly in relation to nonreward variables (Matsumoto and Hikosaka, 2009; Barter et al., 2015; Lerner et al., 2015; Howe and Dombeck, 2016; Menegas et al., 2018; de Jong et al., 2019; Engelhard et al., 2019; Yuan et al., 2019). For instance, dopamine neurons in the medial VTA signal cue-related and outcome-related information, whereas dopamine neurons located in the lateral VTA and SNc appear to encode motor parameters in reward-guided tasks (Howe and Dombeck, 2016; Engelhard et al., 2019). Moreover, these different

Received Jan. 3, 2023; revised Apr. 6, 2023; accepted Apr. 14, 2023.

Author contributions: K.M.F., P.H.J., and R.K. designed research; K.M.F., H.J.P., and R.K. performed research; K.M.F., H.J.P., and R.K. analyzed data; K.M.F., P.H.J., and R.K. wrote the paper.

This research was supported by Department of Health and Human Services | National Institutes of Health | National Institute on Drug Abuse Grant R01-DA-035943.

The authors declare no competing financial interests.

Correspondence should be addressed to Ronald Keiflin at rkeiflin@ucsb.edu or Patricia H. Janak atpatricia.janak@jhu.edu.

https://doi.org/10.1523/JNEUROSCI.0028-23.2023

dopamine cell groups (VTA and SNc) have largely dissociable projection targets, with VTA dopamine (VTA_{DA}) neurons projecting predominantly to the limbic ventromedial striatum, and SNc dopamine (Snc_{DA}) neurons projecting predominantly to the sensorimotor dorsolateral striatum (Björklund and Dunnett, 2007; Ikemoto, 2007; Saunders et al., 2018). The dissociable response profiles and the relatively segregated anatomic targets of VTA and SNc dopamine neurons strongly suggest a functional dissociation between these neuronal populations. In line with this idea, cues paired with optogenetic activation of either VTA or SNc dopamine neuron acquire qualitatively different motivational properties (selective approach vs general locomotion, respectively; Saunders et al., 2018).

Importantly, this regional specialization of dopamine neurons was evident only in Pavlovian conditioning preparations, in which subjects can anticipate—but not control—the delivery of optogenetic dopamine stimulation. This contrasts with the seemingly uniform role for dopamine neurons in self-stimulation preparations, in which the activation of dopamine neurons is contingent on an instrumental response. Indeed, rats will avidly press a lever if this results in the activation of their dopamine neurons, regardless of whether stimulation is delivered to the VTA or SNc (Rossi et al., 2013; Ilango et al., 2014; Saunders et al., 2018; Keiflin et al., 2019).

Why is there strong evidence for functional specialization of VTA and SNc dopamine neurons in Pavlovian but not instrumental reinforcement learning? Although it may be the case that VTA and SNc dopamine neurons contribute uniformly and undistinguishably to instrumental reinforcement, an alternative and more likely hypothesis is that this functional homology is only apparent in reduced or constrained scenarios. Although activation of VTA or SNc dopamine neurons favors the repetition of an instrumental response, the underlying motivational processes engaged by these two neural populations might differ. The purpose of this study was to test this hypothesis of a functional heterogeneity of VTA and SNc dopamine neurons in instrumental reinforcement. For this purpose, rats were trained to press a lever for optogenetic stimulation of VTA or SNc dopamine neurons; these rats were then subjected to different behavioral assays and manipulations designed to probe the nature and content of the processes governing their instrumental response.

Materials and Methods

Animals and surgeries

Male and female transgenic rats expressing Cre recombinase under control of the tyrosine hydroxylase promoter (Th::cre rats) were used in these studies. Rats were individually housed under a 12 h light/dark cycle and had unlimited access to food and water except during testing. The majority of the experiments were conducted during the light cycle. All experimental procedures were conducted in accordance with Johns Hopkins University Institutional Animal Care and Use Committees and the US National Institutes of Health guidelines. Males and females were distributed as evenly as possible across groups. No significant effects of sex were found; therefore, data for males and females were collapsed. Rats (weight: males, >300 g; females, >225 g) were anesthetized with isoflurane (induction, 5%; maintenance, 1–2%) and received unilateral infusions of AAV5-EF1a-DIO-ChR2-eYFP (titer, $\sim 10^{12}$) into the VTA or the SNc under stereotaxic guidance. VTA was targeted as follows: anteroposterior (AP), −5.4 and −6.2 mm from bregma; mediolateral (ML), ±0.7 mm from midline; dorsoventral (DV), −8.5 and −7.5 mm from skull). SNc was targeted at AP −5.0 and −5.8 mm from bregma; at ML ±2.4 mm from midline; and at DV −8.0 and −7.0 mm from skull. This resulted in four injection sites for each rat. At each injection site, 0.5 μl of virus was delivered at the rate of 0.1 μl/min. Injectors were left in place for 10 min following the infusion to allow for diffusion. Immediately following viral infusions, optic fibers (core diameter, 300 μm; numeral aperture, 0.37) aimed at VTA (AP, −5.8 mm from bregma; ML, ±0.7 mm from midline; DV, −7.5 mm from skull) or SNc (AP, −5.4 mm from bregma; ML, ±2.4 mm from midline; DV, −7.2 mm from skull) were implanted. Behavioral training started 3–4 weeks postsurgery to allow for gene expression.

Optical activation

Rats were tethered to optical patch cords (200 µm) connected to a 473 nm blue laser diode (Opto Engine) through a rotary joint (Doric Lenses). Fiber optic implants and patch cords were constructed in the laboratory and were equipped with a custom lock-in mechanism that ensured secure tethering during long behavioral sessions. An individual stimulation event consisted of a 2 s train of light pulses (20 Hz, 40 pulses, 5 ms pulse duration). Unless specified otherwise, the laser output during optogenetic stimulation was 24 mW, resulting in an irradiance of ~8.5 $mW/mm^2/s$ at the tip of the intracranial fiber (corrected for duty cycle). Light power was verified before and after every behavioral session.

Intracranial self-stimulation

Intracranial self-stimulation (ICSS) sessions were conducted in 12 identical sound-attenuated operant chambers (Med Associates). Operant chambers were fitted with two retractable levers on the front panel and one nose-poke operandum on the back panel (obstructed in all experiments with the exception of the heterogeneous instrumental chain experiment). During ICSS sessions, a response at the active lever (position counterbalanced) resulted in the delivery of a 2 s train of light pulses. Inactive lever responses, and active lever responses occurring during the 2 s light train, were recorded but had no consequence. Ventilation fans provided background noise of 65 dB, and a red houselight provided diffuse background illumination during the ICSS sessions.

Initial acquisition. Before being assigned to the different experiments, all rats were initially trained to acquire an instrumental ICSS response. A minimum of 100 active lever presses per hour, for three consecutive sessions, constituted the criterion for successful acquisition (range, 3–11 sessions). The experiments described here were conducted in four different replication cohorts. The initial acquisition of ICSS was conducted under experimental conditions that differed slightly between cohorts and as a function of the different experiments [session time in the day/light cycle, session duration (1–6 h), and the presence or absence of an inactive lever]; therefore, acquisition data were not analyzed. Following initial acquisition, all subsequent experiments were conducted in identical conditions for all cohorts. These differences in procedures during initial ICSS acquisition had no consequences on later behavioral outcomes (no effect of initial training protocol).

Experiment 1: ICSS response patterns. Rats (VTA, $n = 20$; SNc $n = 20$) were given three to five daily 4 h sessions of ICSS. Response patterns were analyzed for the last session completed. In addition, over a series of sessions, a subset of VTA rats ($n = 6$) were tested for the ability of manipulations of light intensity to reproduce the pattern of responding resulting from ICSS of SNc_{DA} neurons. In these sessions, we reduced the light intensity from the fiber over a range of intensities from 24 mW (the intensity used for all experiments otherwise described here) to 1 mW, in addition to testing with no light delivery. These tests were conducted in descending order from the highest (24 mW) to the lowest (0 mW) light intensities. Two or three 4 h consecutive sessions were conducted at each laser intensity; data from the last session are reported.

Experiment 2: forced time-out. To determine the influence of forced time-outs on instrumental responding for VTA_{DA} of SNc_{DA} neurons stimulation, a subset of rats included in experiment 1 (VTA, $n = 15$; SNc, $n = 16$) went on to complete the following experiment. During a single daily 2 h ICSS session, each response on the active lever resulted in an optical stimulation (2 s) and a retraction of both levers (active and inactive) for a period of 12 s. At the end of this delay, both levers were extended and remained extended until the subsequent stimulation.

Experiment 3: lever relocation. Rats (VTA, $n = 10$; SNc, $n = 8$) were trained to lever press for optical stimulation of dopamine neurons in the presence of a single (active) lever, in daily 2 h sessions. To facilitate the detection of the lever, a discrete cue light (28 V) located immediately above the lever remained on during all sessions. After five to six training

sessions, all rats were tested in two probe sessions during which no stimulation was delivered. In one probe session, the lever remained in its usual location. In the other probe session, the panel containing the lever was reconfigured to raise the lever (and the light above the lever) by 8 cm. Animals were retrained for three sessions between the two probe tests (with the lever in its standard position), and the order of testing was counterbalanced within groups.

Experiment 4: cued progressive ratio. Rats were trained to self-stimulate VTA_{DA} ($n = 6$) or SNc_{DA} ($n = 6$) neurons, and each stimulation coincided with the presentation of a brief audiovisual cue (white noise and chamber illumination, 0.4 s). To increase the relevance of the cue and to promote the association between this cue and the optoactivation of dopamine neurons, the response requirement was gradually increased over the course of 8 d, from fixed ratio 1 (FR1) to FR5. We noticed that several rats in the SNc_{DA} group had difficulty maintaining responding under these higher response ratios; therefore, the response requirement was brought back to Random Ration (RR) 1.5 or RR2 for the remaining training sessions. To ensure an equal number of cue–stimulation pairings in both groups, we imposed a maximal number of stimulations per session, which all animals completed (maximum number of stimulations, 200; with the exception of FR3 and FR5 sessions for which the maximum number of stimulations was reduced to 100 and 20, respectively). The conditioned incentive properties acquired by the audiovisual cue were then assessed in four progressive ratio (PR) probe tests. Under the PR schedule, the number of responses required to earn a stimulation was increased after each stimulation according to the following formula (Richardson and Roberts, 1996):

$$\text{Response ratio} = [5e^{(\text{stimulations} * 0.2)}] - 5.$$

This produced the following response requirement schedule: 1-2-4-6-9-12-15-20-25-32-40-50-62.

In two PR test sessions (sessions 1 and 3), the audiovisual cue continued to be presented on every stimulation (as in training). In the other two PR test sessions (sessions 2 and 4), the audiovisual cue was contingent on every response on the active lever (responses produced during cue presentation did not prolong the cue). PR probe sessions were separated by 2 d of retraining under the RR1.5 or RR2 schedule described above.

Experiment 5: heterogeneous instrumental chain. Rats were initially trained, in 4 h sessions, to press an active "taking" lever to obtain an optoactivation of either VTA_{DA} ($n = 13$) or SNc_{DA} ($n = 14$) neurons. A second, inactive, lever was present at this stage, but had no programmed consequence. After five or six sessions under this reinforcement schedule, rats were required to perform a sequence of two instrumental actions to obtain an optogenetic stimulation. Specifically, rats were required to press a "seeking" lever (corresponding to the previously inactive lever) to gain access to the taking lever. A response to the seeking caused the insertion of the taking lever. A press on the taking level would then produce the optogenetic stimulation of VTA_{DA} or SNc_{DA} neurons, and the retraction of the taking lever. After three sessions under this reinforcement schedule, rats were required to perform a sequence of three instrumental actions to obtain an optostimulation. Specifically, rats were required to perform a nose-poke (on the back panel, opposite to the levers), then press the seeking lever, and finally press the taking lever to obtain an optostimulation. An LED light located inside the nose-poke operandum signaled that this operandum was available for responding. A nose-poke response caused the termination of this LED light and the insertion of the seeking lever. A response of the seeking lever caused the insertion of the taking lever. A press on the taking lever would then produce the optostimulation as well as the retraction of all levers and the illumination of the nose-poke LED light. Rats were trained on this schedule for three sessions.

Some rats completed more than one experiment (Table 1, allocation of rats to the different experiments).

Histology

Animals were anesthetized with pentobarbital and perfused intracardially with 0.9% saline followed by 4% paraformaldehyde. Brains were extracted, postfixed in 4% paraformaldehyde for 24 h, and cryoprotected in 30% sucrose for >48 h. Brains were then frozen on dry ice and sectioned at 40 mm on a cryostat. Coronal slices were collected onto glass slides and coverslipped with Vectashield mounting medium with DAPI. Fiber tip position and eYFP-CHR2 virus expression were verified under a fluorescence microscope (Zeiss Microscopy).

Statistical analysis

Statistical analyses were performed using SPSS Statistics package (version 25.0.0.1; IBM SPSS). Graphs were generated using GraphPad Prism 9 and MATLAB R2016a (MathWorks). For each dataset, Lilliefors and Brown–Forsythe tests were run to test for normality and equal variance, respectively. When appropriate, parametric tests were conducted and consisted generally of mixed-design repeated-measures ANOVAs with Group (VTA_{DA} or SNc_{DA}) as a between-subject factor and Session as a within-subject factor. *Post hoc* and planned comparisons were carried out with two-tailed Student's t tests. When non-normality and/or unequal variances were observed in our dataset, nonparametric tests were conducted and consisted of Mann–Whitney U test, or Wilcoxon signed-rank z test (for between-subjects or within-subjects comparison, respectively). Significance was assessed against a type I error rate of 0.05. Additionally, for each pairwise or independent comparison, we provide a 95% confidence interval of the difference of means, which was determined by a bias-corrected and accelerated bootstrapped method (free from distributional assumptions) with 5000 resamples. Effect sizes were estimated with the rank-biserial correlation.

Results

To achieve selective control of midbrain dopamine neurons, we injected a Cre-dependent viral vector for the expression of channelrhodopsin 2 (ChR2) into the VTA or SNc of transgenic TH-Cre rats and implanted an optic fiber aimed at those regions (Fig. 1). All behavioral procedures were conducted 3–6 weeks after surgeries. In all experiments described below, rats were initially trained to press one of two levers (designated as active) to obtain a brief optogenetic activation of VTA or SNc dopamine neurons (2 s stimulation, at 20 Hz). We selected a 2 s stimulation as this approximates the known activity of midbrain dopamine neurons during reward receipt across a variety of studies (Day et al., 2007; Roitman et al., 2008; Witten et al., 2011; Cone et al., 2016; Saunders et al., 2018; Ferguson et al., 2020; Grove et al., 2022; van Elzelingen et al., 2022). Rats rapidly acquired reliable and stable self-stimulation responding (within 3–11 sessions; data not shown). Responding on the inactive lever was low for both groups (mean ± SEM; VTA, 0.844 ± 0.236 responses/h after initial ICSS acquisition; SNc, 1.219 ± 0.312 responses/h after initial ICSS acquisition; $U = 235.5$, $p = 0.276$; 95% CI, −0.342, 1.135; effect size, 0.182) and remained negligible throughout the different manipulations. Therefore, for simplicity, we will only present active-lever presses.

Self-stimulation of VTA or SNc dopamine neurons generates different patterns of instrumental responding

The response-contingent activation of VTA or SNc dopamine neurons resulted in striking differences in instrumental response patterns (Fig. 2). Subjects self-stimulating VTA_{DA} neurons displayed high rates of responding throughout the 4 h session. In contrast, responding for SNc_{DA} neuron activation was characterized by bouts of vigorous responding interrupted by long pauses. This resulted in significant group differences in the total number of operant responses ($U = 16$, $p < 0.001$; 95% CI, 5988.18, 12 595.94; effect size, 0.920) and the average inter-response intervals (IRIs; $U = 16.00$, $p < 0.001$; 95% CI, 3.78, 6.92; effect size, 0.920). To assess the regularity of responding within a session, we calculated for each subject the coefficient of variation

Table 1. Subjects' allocation to the different experiments

	Response Pattern	Imposed time-outs	Raised lever	Progressive ratio	Instrumental chain
VTA 01	X	X			X
VTA 02	X	X			X
VTA 03	X	X			X
VTA 04	X	X			X
VTA 05	X	X			
VTA 06	X	X			
VTA 07	X	X			
VTA 08	X	X			
VTA 09	X	X			
VTA 10	X	X			
VTA 11	X	X			
VTA 12	X	X			
VTA 13	X	X			
VTA 14	X	X			
VTA 15	X	X			
VTA 16	X				X
VTA 17	X				X
VTA 18	X				X
VTA 19	X				X
VTA 20	X				X
VTA 21			X		
VTA 22			X		
VTA 23			X		
VTA 24			X		
VTA 25			X		
VTA 26			X		
VTA 27			X		
VTA 28			X		
VTA 29			X		
VTA 30			X		
VTA 31				X	
VTA 32				X	
VTA 33				X	
VTA 34				X	
VTA 35				X	
VTA 36				X	
VTA 37					X
VTA 38					X
VTA 39					X
VTA 40					X
SNc 01	X	X			X
SNc 02	X	X			X
SNc 03	X	X			X
SNc 04	X	X			X
SNc 05	X	X			X
SNc 06	X	X			X
SNc 07	X	X			
SNc 08	X	X			
SNc 09	X	X			
SNc 10	X	X			
SNc 11	X	X			
SNc 12	X	X			
SNc 13	X	X			
SNc 14	X	X			
SNc 15	X	X			
SNc 16	X	X			
SNc 17	X				X
SNc 18	X				X
SNc 19	X				X
SNc 20	X				X
SNc 21			X		
SNc 22			X		
SNc 23			X		
SNc 24			X		
SNc 25			X		
SNc 26			X		
SNc 27			X		
SNc 28			X		
SNc 29				X	
SNc 30				X	
SNc 31				X	
SNc 32				X	
SNc 33				X	
SNc 34				X	
SNc 35					X
SNc 36					X
SNc 37					X
SNc 38					X

of IRIs. Subjects self-stimulating SNc_{DA} neurons displayed a higher coefficient of variation ($U = 21.00$, $p < 0.001$; 95% CI, 4.87, 9.44; effect size, 0.895), consistent with the irregular pattern of responding in this group. Despite significant differences between VTA and SNc groups in average IRIs, the relative distribution of IRIs appears remarkably similar in both groups, with the vast majority (>90%) of responses occurring in rapid succession (IRI, ≤4 s). Where VTA and SNc groups differ most profoundly is in the duration of their instrumental pauses (pauses were arbitrarily defined as a period of ≥20 s that separates two instrumental responses). These pauses were equally frequent in both groups ($t = 0.698$, $p = 0.490$; 95% CI, −13.79, 28.37; effect size, 0.335), but lasted much longer in subjects self-stimulating SNc_{DA} neurons ($U = 9.00$, $p < 0.001$; 95% CI, 144.36, 264.51; effect size, 0.955). Analysis focused on periods of peak responding in a session confirmed that when actively engaged in responding both VTA and SNc groups responded at the same rate ($U = 183.00$, $p = 0.655$; 95% CI, −0.35, 0.23; effect size, 0.085).

To determine whether these different response patterns could reflect different reward intensities induced by VTA versus SNc dopamine neuron stimulation, we systematically reduced the intensity of the optostimulation (by reducing the laser power) in a subset of VTA_{DA} self-stimulating rats ($n = 6$) following the acquisition of stable responding. While reductions in VTA light intensity could replicate some aspects of SNc_{DA} self-stimulation, we were unable to fully replicate the behavioral pattern emitted by SNc_{DA} rats in VTA_{DA} rats (Fig. 3). Importantly, we never observed at any light intensity the long pauses in responding that were characteristic of SNc_{DA} rats. Note, however, that the failure to replicate these long instrumental pauses may be because of the prior training of VTA_{DA} rats at the higher laser intensity. These results suggest that, rather than producing different intensities of reinforcement, the activation of VTA_{DA} or SNc_{DA} neurons engages different reinforcement processes.

Imposed time-outs abolish instrumental responding for SNc, but not VTA, dopamine neurons self-stimulation

The irregular pattern of responding observed in the SNc group suggests that, unlike subjects in the VTA_{DA} group, animals self-stimulating SNc_{DA} neurons might lack the motivation to approach the active lever and initiate bouts of responding (this is despite the fact that both groups respond avidly for dopamine self-stimulation once a bout has been initiated). However, given

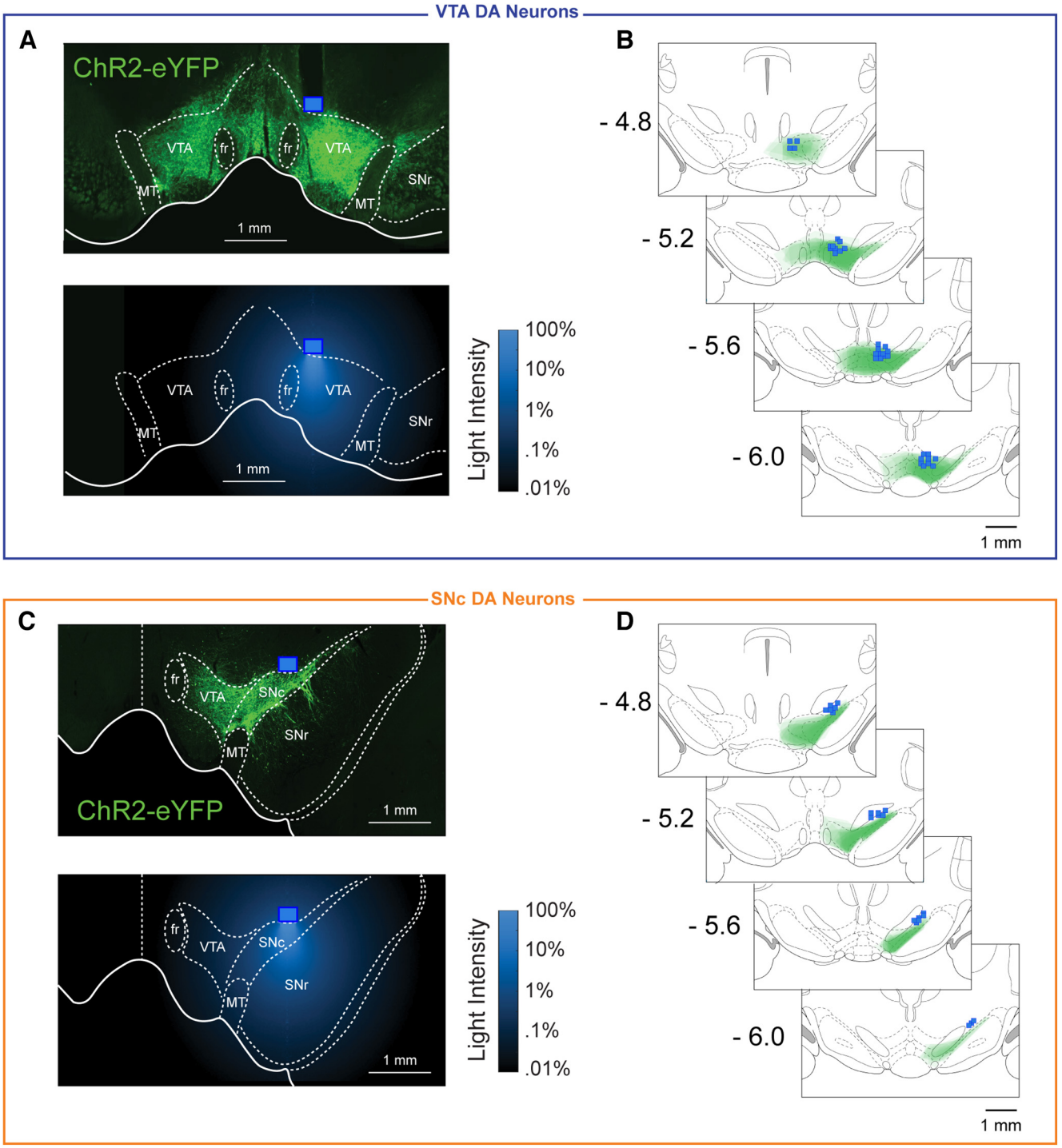

Figure 1. Histologic verification of ChR2 expression in VTA and SNc DA neurons. ***A***, Representative expression of ChR2-eYFP in the VTA of TH-Cre rats and an accompanying estimate of relative laser intensity and spread from the optic fiber tip (Stujenske et al., 2015). ***B***, Reconstruction of expression and optic fiber placements for all animals in the VTA group. Expression is indicated by overlays of total spread for each rat. Blue squares denote the ventral extremity of fiber implants. ***C***, Same as ***A*** but for a representative rat in the SNc group. ***D***, Same as ***B*** but for TH-Cre rats in the SNc group.

the established role of the nigrostriatal dopamine pathway in motor functions (Dodson et al., 2016; Howe and Dombeck, 2016; da Silva et al., 2018), an alternative explanation for the irregular and interrupted pattern of responding in the SNc_{DA} group is that the activation of SNc_{DA} neurons produces motor effects that are incompatible with sustained high rates of instrumental responding (Grilly, 1977). The following experiment was intended to tease apart the contribution of these two potential mechanisms (motivation to initiate responding vs competing motor effects).

After initial self-stimulation training (as described above), rats were tested in a session in which each press on the active lever resulted in the optogenetic activation of VTA ($n = 15$) or SNc ($n = 16$) dopamine neurons and was immediately followed by the retraction of both active and inactive levers for a duration of 12 s. After this imposed time-out period, both levers were again

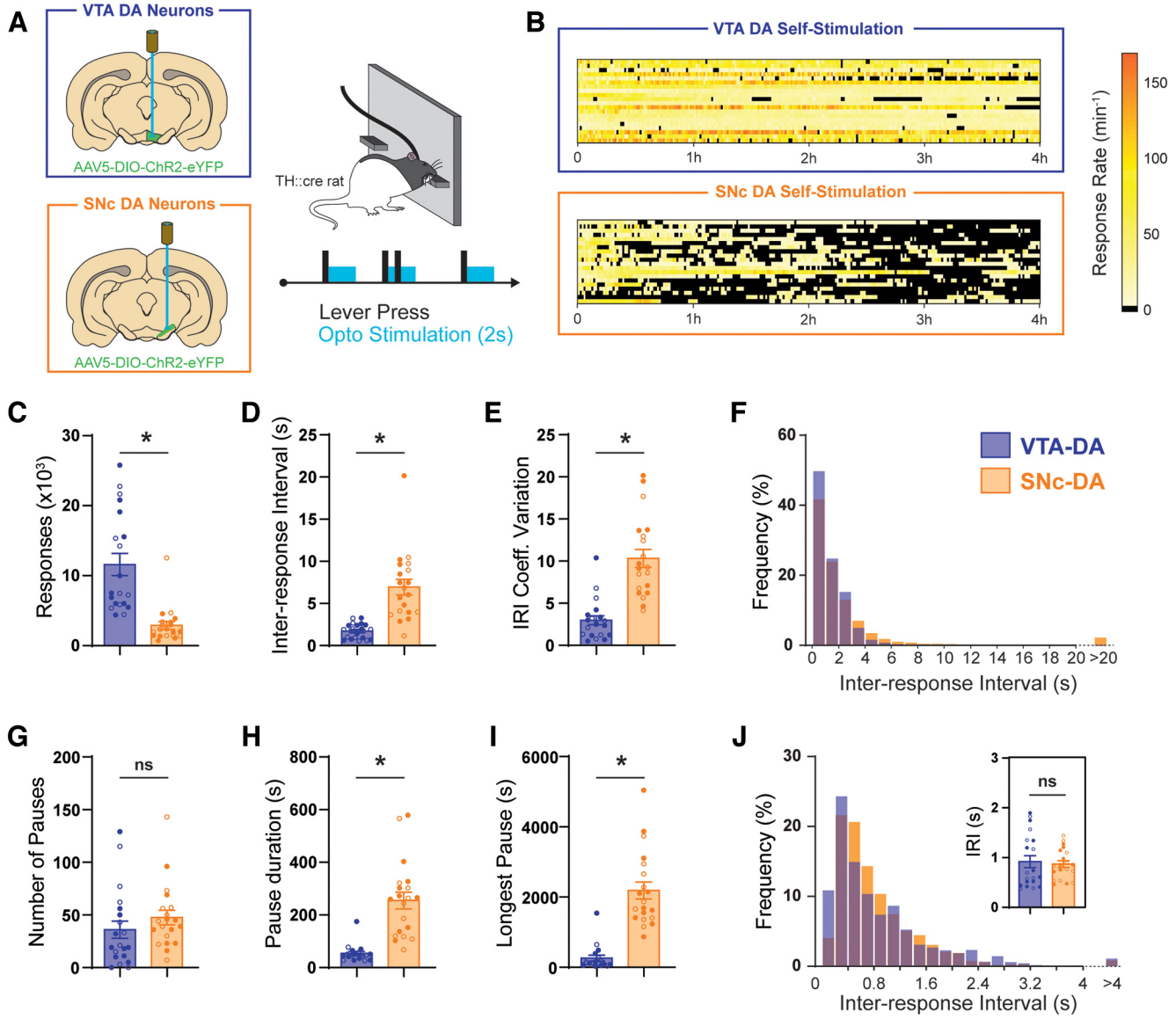

Figure 2. Self-stimulation of VTA or SNc dopamine neurons produces different patterns of operant responses. ***A***, TH::Cre rats were made to express ChR2 in either VTA ($n = 20$) or SNc ($n = 20$) dopamine neurons. Responding on the active lever resulted in a 2 s optoactivation of the targeted dopamine neurons. ***B***, Heat maps of the rate of operant responding throughout the 4 h sessions, in animals self-stimulating VTA_{DA} or SNc_{DA} neurons. Each line represents a different animal. ***C***, Total responses on the active lever. ***D***, Average IRI. ***E***, Coefficient of variation of IRI. ***F***, Frequency distribution of IRIs (as a percentage of total responses). ***G***, Number of within-session pauses in operant responding (pause defined as an IRI >20 s). ***H***, Average pause duration. ***I***, Longest pause in a session. ***J***, Frequency distribution of IRIs during a 4 min peak responding period (*$p < 0.001$, Mann–Whitney *U* tests; ns: not significant). Filled symbols, males; empty symbols, females. Error bars indicate the SEM.

extended, and rats could press the active lever for a subsequent stimulation (Fig. 4). This reinforcement schedule limits the number of stimulations that can be obtained in a session; it also allows for potential stimulation-induced motor effects to dissipate before the next opportunity to respond (stimulation lasts for 2 s but the time-out is 12 s). We reasoned that if motor effects are responsible for the reduced responding in the SNc group, then the imposed time-outs should mitigate these motor effects and tend to equalize responding between the two groups. On the other hand, if reduced responding in the SNc group is because of a reduced motivation to initiate responding after a pause, then the imposed time-outs should further reduce responding in that group, as every lever press now requires subjects to initiate responding by approaching or reapproaching the lever (i.e., continuous high-frequency presses are no longer possible).

To facilitate comparison between continuous reinforcement and imposed time-out sessions, we calculated the response rate during the period when the levers were present (when a lever press response could actually be produced and result in stimulation). The imposed time-outs significantly decreased response rate in animals self-stimulating SNc dopamine neurons ($z = -3.516$, $p < 0.001$; 95% CI, 8.890; 20.167; effect size, 1.00), but not in animals self-stimulating VTA dopamine neurons ($z = -1.477$; $p = 0.151$; 95% CI, −0.980, 15.367; effect size, 0.333). To directly compare the two groups, we calculated for each subject a suppression ratio, defined here simply as the response rate during the imposed time-out session divided by the response rate during continuous reinforcement. We found that, compared with animals self-stimulating VTA dopamine neurons, animals self-stimulating SNc dopamine neurons were more sensitive to imposed time-outs ($U = 0.000$; $p < 0.001$; 95% CI, 0.726, 1.060;

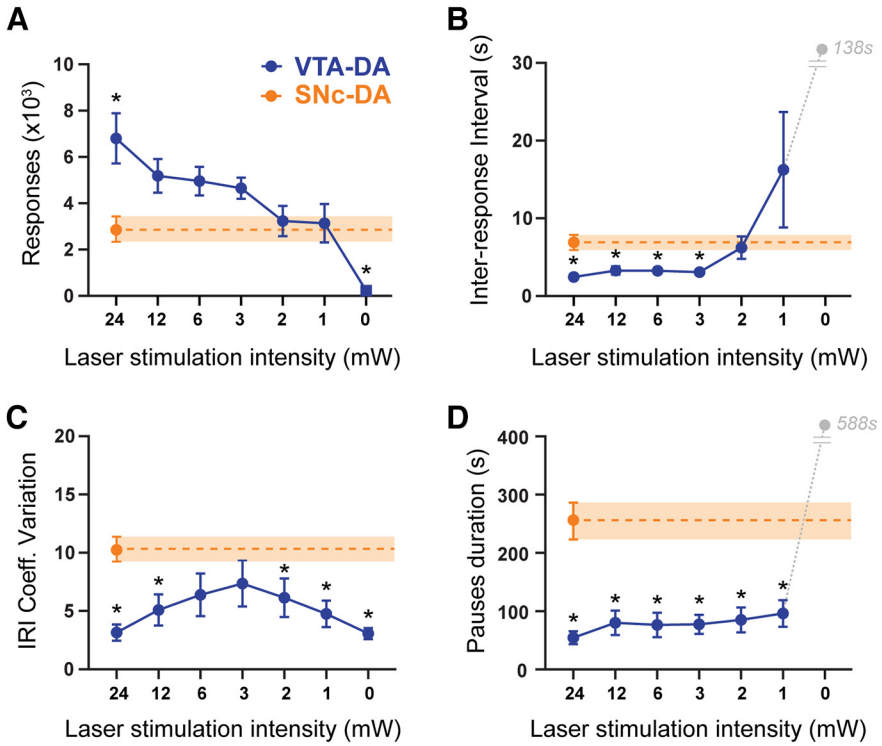

Figure 3. Operant responding for different intensities of VTA_{DA} neuron stimulation. ***A***, Total responses per session, ***B***, IRI. ***C***, Coefficient of variation of IRIs within a session. ***D***, Average pause duration. Reducing VTA-DA stimulation intensity reduced the number of responses and increased the average IRI, but those lower stimulation intensities failed to reproduce the irregular pattern or long pauses in responding that characterize SNc_{DA} self-stimulation. Error bars and error bands indicate the SEM. Orange line and shading represent the average ± SEM for the SNc group at the maximum intensity of 24 mW. *$p < 0.05$, significant difference from SNc_{DA} 24 mW, bootstrapped *t* tests.

effect size, 1.00). This strongly suggests that, compared with animals in the VTA group, animals self-stimulating SNc dopamine neurons express reduced motivation to reengage with the lever and reinitiate responding after a pause, whether that pause is experimentally imposed or spontaneous (self-imposed).

Topological changes in response requirement abolish instrumental responding for SNc, but not VTA, dopamine neuron self-stimulation

Our results thus far suggest that self-stimulation of dopamine neurons in the VTA or SNc engages qualitatively different instrumental processes. While subjects in both groups are capable of highly stereotyped responding within a bout, SNc subjects appear deficient in their ability to return to the lever and reinitiate responding after a pause. An important distinction between within-bout responding and bout initiation is that only the latter requires a flexible approach strategy, since for every bout initiation (and depending on the initial position and location of the animal in the chamber) a new set of actions is required to approach and reach the lever (Nicola, 2010). In the following experiment, we decided to further interrogate a component of flexible approach in instrumental responding by imposing environmental constraints that required subjects to perform a new—topologically different—response to reach and press the lever.

Rats initially trained to press a lever to self-stimulate VTA ($n = 10$) or SNc ($n = 8$) dopamine neurons (average ± SEM response rate: VTA_{DA} = 36.4 ± 7.29 responses/min; SNC_{DA} = 9.95 ± 2.31 responses/min; $U = 8.0$; $p = 0.005$) were tested in the following two probe sessions: in one session, the lever remained in its standard position, and in another session the lever was raised by 8 cm (order counterbalanced, 3 d of retraining between probe sessions). Importantly, optogenetic stimulation was not delivered in these probe tests to avoid the confounding effects of within-session reinforcement of the new response. The raised lever was still within reach, but deflecting the lever now required a different set of motor commands, resulting in a different body posture (crouching in the standard lever position vs rearing in the elevated lever position; Fig. 5). Throughout this experiment, only the active lever was presented to prevent the potential transfer of responding between levers. Moreover, to facilitate the detection of the lever, a stimulus light located above the lever was continuously illuminated (during training and probe tests).

Performance when rats had to rear to reach the raised lever was compared with their responding in a session in which the lever was in its usual position. In neither session did lever pressing lead to stimulation. In the absence of reinforcement, instrumental responding during probe sessions extinguished extremely rapidly for both groups (a phenomenon commonly observed in electrical or optical self-stimulation preparations; Olds and Milner, 1954; Witten et al., 2011). A two-way mixed ANOVA (Group × Session) found no main Group effect ($F_{(1,16)} = 4.093$, $p = 0.060$). However, the ANOVA revealed a significant Session effect ($F_{(1,16)} = 32.029$, $p < 0.001$) and Group × Session interaction ($F_{(1,16)} = 4.971$, $p = 0.04$). Planned *post hoc* comparisons indicated that while SNc and VTA groups did not differ in a standard extinction session ($t = 0.179$; $p = 0.859$; 95% CI, −23.00, 19.10; effect size, 0.075), the relocation of the lever induced a reduction in responding that was much more pronounced in the SNc group, resulting in a significant difference between these groups ($t = 2.974$; $p = 0.006$; 95% CI, −46.726, −15.146; effect size, 0.824). This indicates that, unlike subjects in the SNc group, subjects in the VTA group were able to improvise a new set of actions to seek out and reach the relocated lever despite that lever not resulting in reinforcement.

Cues paired with VTA, but not SNc, dopamine neuron stimulation increase responding in progressive ratios

Our results thus far indicate that, compared with animals in the SNc group, animals in the VTA group express a higher propensity to approach the active lever operandum. This suggests that in the VTA group, the circumstances (spatial location and/or environmental stimuli) surrounding optogenetic stimulation have acquired some incentive properties that compel subjects to approach the active lever and initiate responding. In addition to their ability to motivate approach, another defining property of incentive stimuli is that animals will work to obtain those stimuli, even in the absence of the primary reward they signal. Therefore, in this next experiment, we decided to formally test the incentive properties acquired by phasic stimuli paired with VTA or SNc dopamine neuron self-stimulation by testing their

ability to maintain instrumental responding in the (near) absence of actual optogenetic stimulation.

Rats were trained to self-stimulate VTA ($n = 6$) or SNc ($n = 6$) dopamine neurons, and each stimulation coincided with the presentation of a brief audiovisual cue (white noise and chamber illumination, 0.4 s; Fig. 6). To facilitate the association between this cue and the optogenetic stimulation of dopamine neurons, we progressively increased the response requirement (from FR1 to RR2), which reduced the response–stimulation contingency, while maintaining a maximal cue–stimulation contingency—effectively increasing the relevance of the cue. Moreover, to ensure an equal number of cue–stimulation pairings in both groups, we imposed a maximal number of stimulations per session, which all animals completed despite having differing response rates (Fig. 6*B*). Finally, animals were tested in four behavioral sessions in which optogenetic stimulation was delivered according to a PR schedule—a situation in which the vast majority of responses do not result in optogenetic stimulation. In two of those PR test sessions (sessions 1 and 3), the audiovisual cue continued to be presented on every stimulation (as in training). In the other two PR test sessions (sessions 2 and 4), the audiovisual cue was contingent on every response on the active lever. A two-way mixed ANOVA (Group × Session) conducted on the number of responses on the active lever found no main Group effect ($F_{(1,10)} = 3.161$, $p = 0.106$), but did find a significant Session effect ($F_{(3,30)} = 16.937$, $p < 0.001$) and Group × Session interaction ($F_{(3,30)} = 5.284$, $p = 0.013$; Greenhouse–Geisser test corrected). *Post hoc* tests revealed that the VTA and SNc groups did not differ in sessions in which the cue was presented only during receipt of optogenetic stimulation (p values > 0.398). However, the introduction of a response-contingent cue increased responding only for the VTA group (S1 vs S2 and S3 vs S4: p values < 0.010) and not for the SNc group (p values > 0.591), resulting in significant group differences on those session (p values < 0.030). This increase in responding observed in the VTA group resulted in a very modest but significant increase in the number of stimulations obtained by that group (Fig. 6*E*,*F*). The fact that only the VTA group benefited from phasic, response-contingent, optogenetic stimulation-paired cues, indicates that cues paired with the activation of VTA, but not SNc, dopamine neurons acquired incentive properties.

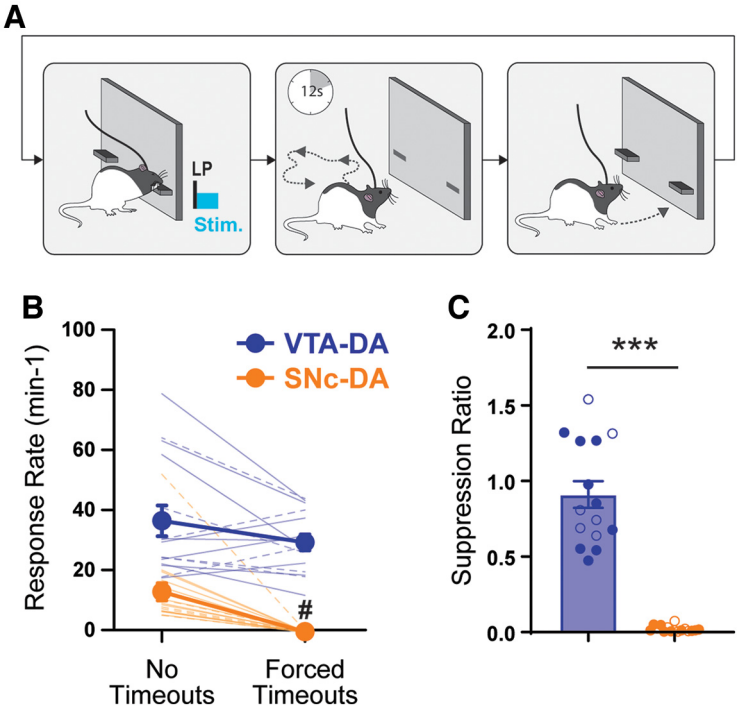

Figure 4. Imposed time-outs abolish responding for SNc_{DA} but not VTA_{DA} neuron stimulation. ***A***, Behavioral paradigm. Each response on the active lever results in optical stimulation of VTA_{DA} ($n = 15$) or SNc_{DA} ($n = 16$) neurons and triggers the retraction of both levers for a period of 12 s (imposed time-out). ***B***, Response rate in the absence or presence of imposed time-outs. ***C***, Suppression ratio induced by imposed time-outs. #$p < 0.001$, Wilcoxon z-test, No Time-outs vs Imposed Time-outs; *$p < 0.001$, Mann–Whitney U test, VTA_{DA} vs SNc_{DA}. Filled symbols and solid lines, males; empty symbols and dashed lines, females. Error bars indicate the SEM.

Activation of VTA, but not SNc, dopamine neurons sustains heterogeneous instrumental sequences

Our results thus far indicate that while the activation of either VTA or SNc dopamine neurons serves as a potent reinforcer of instrumental actions, only in the VTA group does the "state" (location and/or environmental stimuli) associated with dopamine stimulation acquire some incentive value. An evolutionarily advantageous property of a state-value function is that, by signaling when the prospect of reward has increased, it can guide animals through the acquisition of complex instrumental sequences leading to a primary reinforcer (Shahan, 2010; Enquist et al., 2016). In this experiment, we tested to what extent subjects self-stimulating VTA or SNc dopamine neurons could acquire an increasingly complex action sequence to obtain optogenetic stimulation.

Rats were initially trained to press a lever to obtain optogenetic stimulation of either VTA ($n = 13$) or SNc ($n = 14$) dopamine neurons. After five or six sessions (post-initial acquisition) under this reinforcement schedule, rats were then required to perform a sequence of two instrumental actions to obtain an optogenetic stimulation. Specifically, they were required to press one lever (the previously inactive lever) to gain access to another lever (the previously active lever) that they could then press to obtain the stimulation (i.e., a seeking–taking reinforcement schedule). After three sessions under this reinforcement schedule, the required instrumental sequence was further extended (Fig. 7). To obtain an optogenetic stimulation, rats had to perform a sequence of three instrumental actions, starting with a nose-poke, then a press on a first (seeking) lever, and finally a press on a second (taking) lever.

In both groups, the transition from a single action to a two-action sequence induced a sharp decrease in the number of stimulations obtained. However, unlike rats in the SNc group, rats in the VTA group rapidly learned the required sequence and by day 3 obtained ~65% of the stimulations earned in baseline. Note that some reduction in the number of stimulations is expected even in subjects having perfectly learned the new sequence, as it takes longer to complete a two-action sequence than a single action. Likewise, when the instrumental requirement increased from a two-action to a three-action sequence, the number of stimulations obtained by VTA rats abruptly decreased, but most rats eventually learned the new sequence, and by day 3 of this schedule VTA rats obtained ~40% of their baseline stimulations

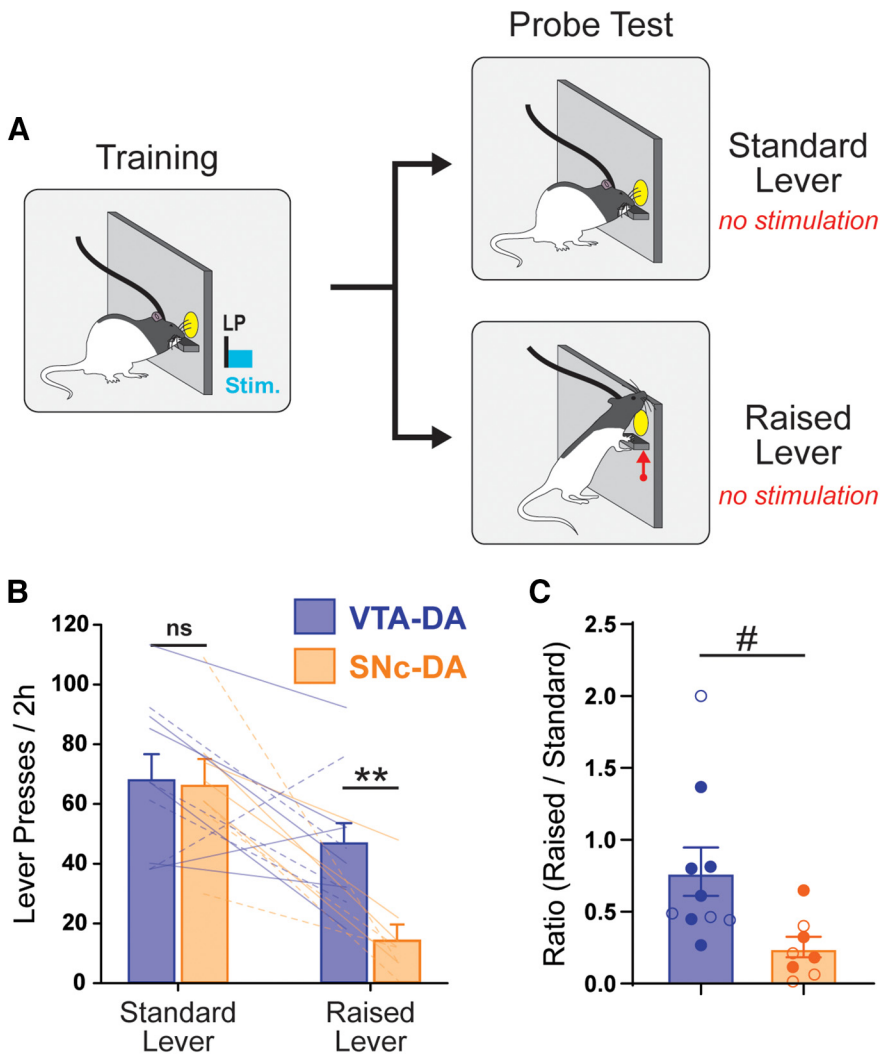

Figure 5. Lever relocation, and modified response requirement, has moderate effects on instrumental responding for VTA_{DA} stimulation, but strongly disrupts instrumental responding for SNc_{DA} stimulation. ***A***, Rats trained to self-stimulate VTA ($n = 10$) or SNc ($n = 8$) dopamine neurons were tested in two nonreinforced probe sessions. For one probe session, the position of the active lever was raised, thereby imposing a new set of motor commands to reach and activate the lever. ***B***, Instrumental responses during probe sessions, with standard or relocated (raised) lever position. ***C***, Suppression ratio induced by lever relocation. ns: not significant $^*p < 0.01$, *t* test; $\#p < 0.01$, Mann–Whitney *U* test; 95% CI, −0.897, −0.213; effect size, 0.8. Filled symbols and solid lines, males; empty symbols and dashed lines, females. Error bars indicate the SEM.

in the one-lever condition. In contrast, the number of stimulations earned by rats in the SNc group remained extremely low.

To compare the two groups, we conducted a two-way mixed ANCOVA (Group × Schedule) on the number of stimulations obtained on the last session of each schedule. To account for the difference in baseline performance between the VTA and SNc groups, the number of stimulations earned at baseline was used as a covariate. This analysis revealed a main effect of Group ($F_{(1,24)} = 18.006$, $p < 0.001$), and Group × Schedule interaction ($F_{(2,48)} = 13.943$, $p < 0.001$), confirming the different impact of the increasing action sequence requirement on VTA_{DA} and SNc_{DA} neuron self-stimulation. Thus, VTA dopamine neuron stimulation can bridge the gap between events and guide learning through a series of actions leading to their activation, whereas SNc dopamine neurons fail to do so.

Discussion

The activation of midbrain dopamine neurons, in the VTA or the adjacent SNc, is a potent reinforcer of instrumental behavior (Rossi et al., 2013; Steinberg et al., 2013; Ilango et al., 2014; Pascoli et al., 2015; Saunders et al., 2018; Keiflin et al., 2019). However, we show here that VTA and SNc dopamine neurons are not functionally equivalent in instrumental reinforcement. Despite superficial similarities in the overt instrumental response, self-stimulation of VTA or SNc dopamine neurons engages different underlying associative structures. Self-stimulation of VTA dopamine neurons was characterized by a high rate of lever pressing, as well as a strong attraction toward the cues/states associated with the stimulation (i.e., the proximity to the active lever or the active lever itself). These incentive, or "motivational-magnet," properties acquired by these cues allowed VTA_{DA} self-stimulating animals to easily reengage in the task following a pause (self-imposed or experimentally imposed time-outs) and to flexibly approach the active lever following its spatial relocation. The incentive properties acquired by the stimuli surrounding VTA_{DA} self-stimulation were also evident in the ability of these stimuli to act as conditioned reinforcers, sustaining instrumental responding in the absence of the primary reinforcer (i.e., in absence of optical simulation). This combination of attracting and reinforcing properties exerted by the environmental stimuli surrounding VTA_{DA} stimulation might have facilitated the acquisition of increasingly complex instrumental sequences for the self-stimulation of VTA_{DA} neurons. For instance, the insertion of the taking lever simultaneously reinforces the previous action responsible for the apparition of this stimulus (i.e., pressing the seeking lever) and attracts the animal toward this new stimulus, with the animal ultimately following a gradient of increasing value, or reward taxis (Karin and Alon, 2022). In contrast, self-stimulation of SNc dopamine neurons was characterized by bouts of a high rate of instrumental responding separated by long pauses during which instrumental responding was completely absent. This response pattern suggests that SNc_{DA} stimulation, while capable of reinforcing an instrumental action, fails to confer incentive properties to the cues/states associated with that stimulation. Consistent with this interpretation, rats self-stimulating SNc_{DA} failed to reengage in the task following an imposed time-out and showed reduced engagement with the active lever following its spatial relocation. Moreover, cues paired with SNc_{DA} stimulation failed to sustain instrumental responding in the absence of the stimulation. Consequently, in the absence of incentive cues to guide them, rats stimulating SNc_{DA} neurons failed to acquire complex instrumental sequences for SNc_{DA} stimulation.

Prior studies showed that the phasic stimulation of SNc_{DA} neurons does not elicit movement initiation, but rather potentiates ongoing movements (Coddington and Dudman, 2018; Saunders et al., 2018; Hamilos et al., 2021). However, another recent study showed that SNc_{DA} stimulation, delivered during

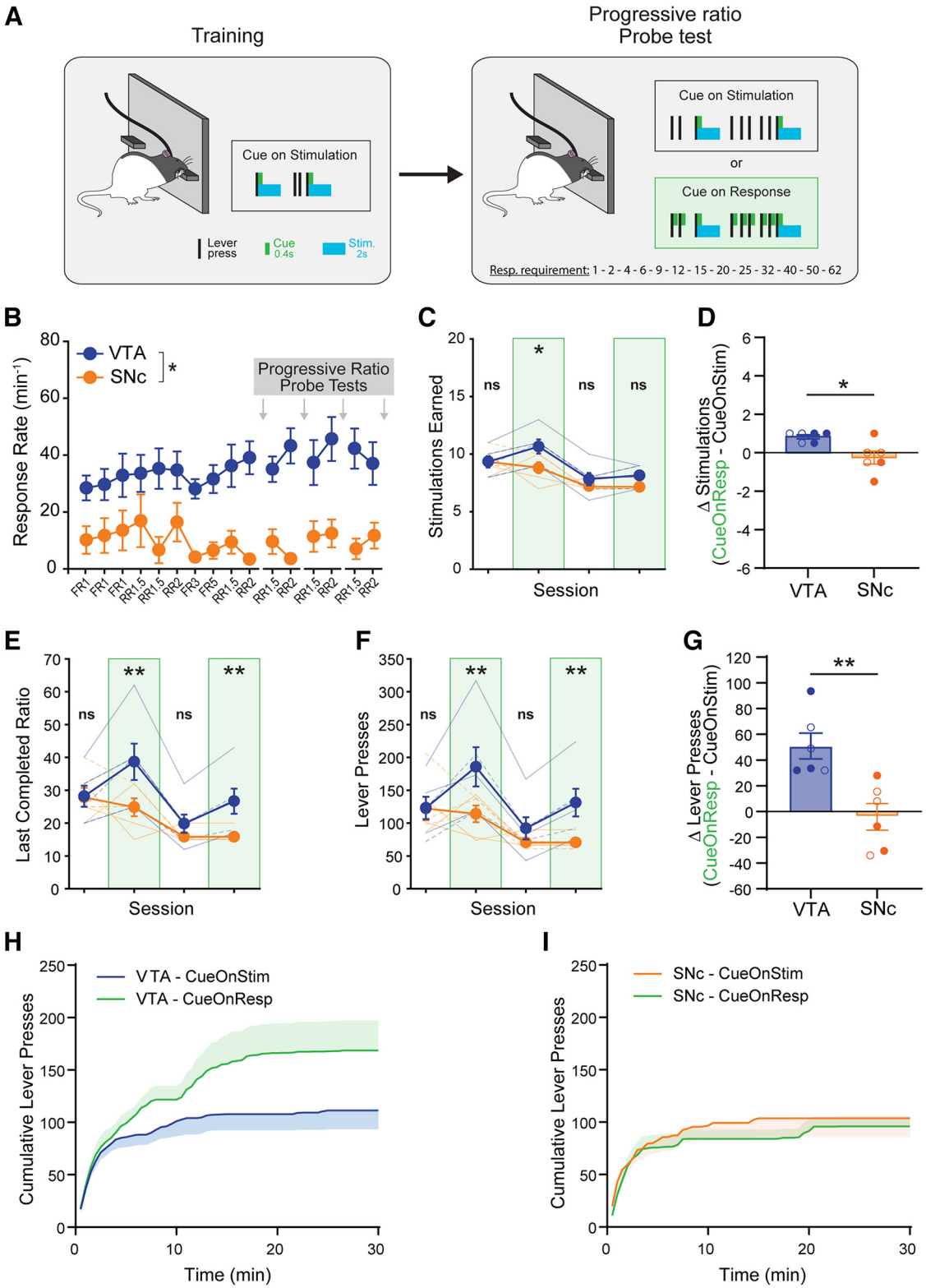

Figure 6. Cues paired with VTA_{DA}, but not SNc_{DA}, neuron stimulation increases responding in progressive ratio. ***A***, Behavioral paradigm. During VTA_{DA} ($n = 6$) or SNc_{DA} ($n = 6$) self-stimulation training, every optostimulation was paired with a brief audiovisual cue. Responding was then tested in progressive ratio sessions, in which the cue continued to accompany every stimulation (cue on stimulation, sessions 1 and 3) or was contingent on every active lever press (cue on response, sessions 2 and 4). ***B***, Rats initially acquired instrumental self-stimulation of VTA_{DA} or SNc_{DA} neurons under a continuous schedule of reinforcement (FR1). To strengthen the association between this cue and the optostimulation, the response requirement was progressively increased (FR1, RR2, FR3, and FR5); thereby reducing the response–stimulation contingency, while maintaining a maximal cue–stimulation contingency. Because several rats in the SNc_{DA} group had difficulty maintaining responding under these higher response ratios, the response requirement was brought back to RR1.5 or RR2 for the remaining training sessions. All rats obtained the maximum number of stimulations on each session, although the VTA_{DA} rats had a higher response rate (*Main Group effect: $F_{(1,10)} = 13.86$, $p = 0.004$). ***C***, ***D***, Number of stimulations earned (***C***) and difference score obtained by subtracting the average stimulations earned in sessions 1 and 3 from the average in sessions 2 and 4 (***D***). ***E***, ***F***, Highest completed ratio (***E***) and total lever presses (***F***) in progressive ratio probe sessions. ***G***, Difference score. ***H***, ***I***, Response time course (cumulative responses) during progressive ratio probe tests for VTA (***H***) and SNc (***I***) groups. *$p < 0.05$, **$p < 0.01$ *t* tests, VTA versus SNc; ns: not significant. Filled symbols and solid lines, males; empty symbols and dashed lines, females. Error bars indicate the SEM.

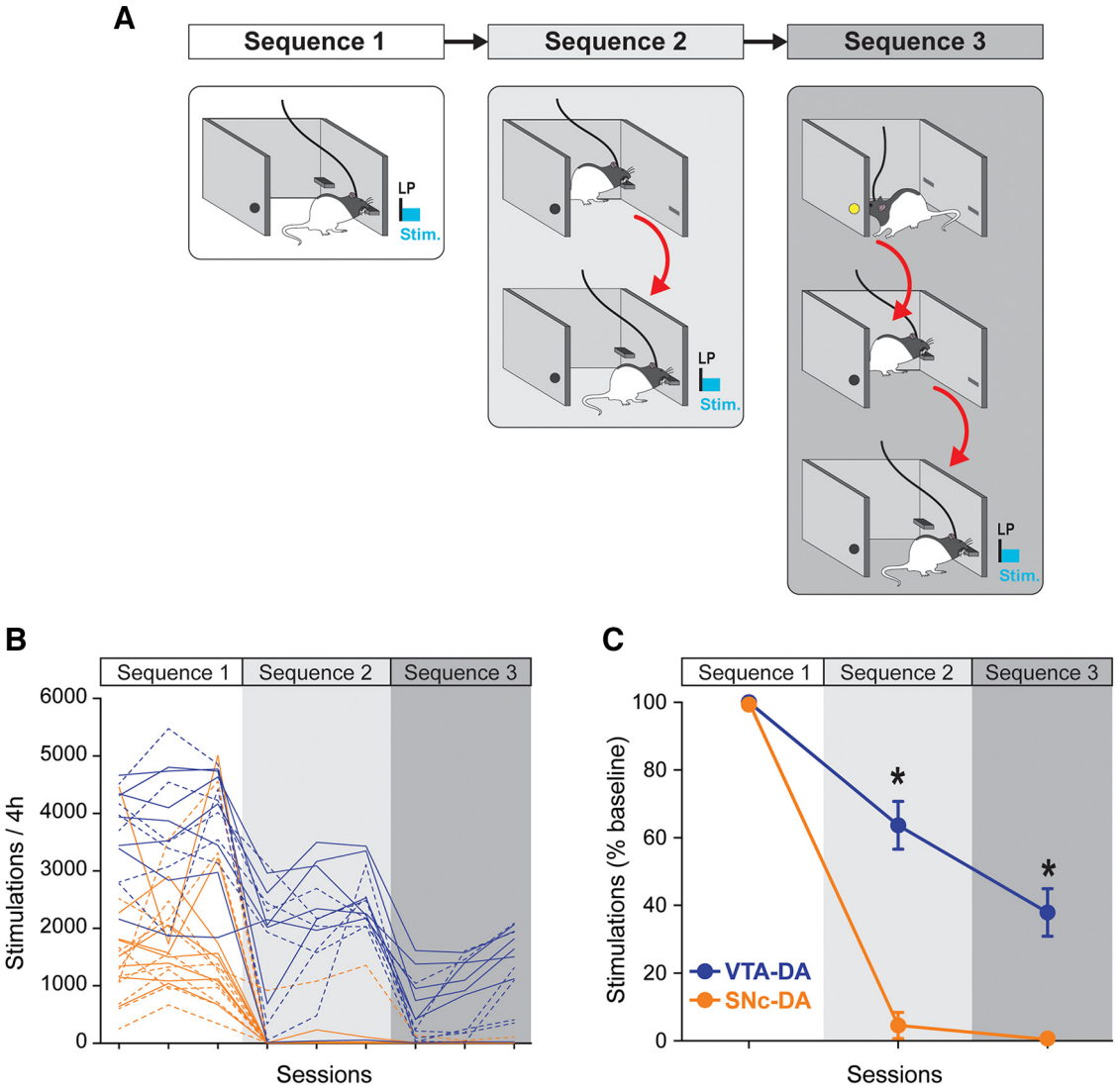

Figure 7. Acquisition of heterogeneous instrumental chain for VTA_{DA}, but not SNc_{DA}, neuron self-stimulation. ***A***, Behavioral paradigm. Rats were initially trained to perform a single instrumental action (press a lever) to obtain an optical stimulation of VTA ($n = 14$) or SNc ($n = 14$) dopamine neurons. The instrumental requirement for self-stimulation was then increased to a sequence of 2, then 3, instrumental actions (see text for details). ***B***, Number of stimulations obtained under the different instrumental requirements. Each line represents an individual subject. ***C***, Average number of stimulations obtained on the final session of each instrumental requirement, expressed as a percentage of baseline (baseline = number of stimulations obtained under sequence 1). *$p < 0.001$, *post hoc* Bonferroni's *t* test. Solid lines, males; dashed lines, females. Error bars indicate the SEM.

fine movements, could also disrupt fine motor kinematics (Bova et al., 2020). Therefore, although we did not observe any evidence for gross motor impairments in SNc_{DA} self-stimulating rats, it remains possible that the stimulation of SNc_{DA} neurons might have subtle motor effects that could warp the precise motor commands for the instrumental response. This might have contributed to the reduced rates of lever pressing observed in those rats and their reduced motivation to initiate responding. Note, however, that SNc_{DA}-stimulating and VTA_{DA}-stimulating rats displayed similar rates of lever pressing during response bouts, when stimulations of DA neurons were most frequent. Instead, it was during instrumental pauses—when no DA stimulation occurred—that VTA and SNc rats differed the most (VTA_{DA} being more likely to re-engage in the task after a pause).

This study indicates that a limiting factor in SNc_{DA} self-stimulation, and the reason for the overall lower responding in that group, is a reduced motivation to initiate responding. This finding might appear at odds with those of previous studies (including work from our group), which reported comparable levels of responding for VTA_{DA} and SNc_{DA} stimulation (Ilango et al., 2014; Saunders et al., 2018; Keiflin et al., 2019). Several factors might explain this discrepancy, including (1) the nature of the operandum, (2) the novelty of the operandum, and (3) the presence or absence of ancillary stimuli. Indeed, in our previous studies, we used nose-pokes as operanda (vs levers in the present study), which generally allow for higher levels of reinforced, but also spontaneous (nonreinforced), responding (Mekarski, 1988; Schindler et al., 1993; Clemens et al., 2010). Moreover, in those same studies, the nose-poke operanda were made available after animals experienced several Pavlovian conditioning sessions in the same chambers. The introduction of this novel operandum in a safe and familiar environment is likely to increase spontaneous exploration and interaction (Ennaceur and Delacour, 1988). Finally, unlike the present study, other studies used ancillary stimuli (during initial training) in the form of brief visual stimuli accompanying VTA_{DA} or SNc_{DA} neuron stimulation (Ilango et al., 2014). Such response-

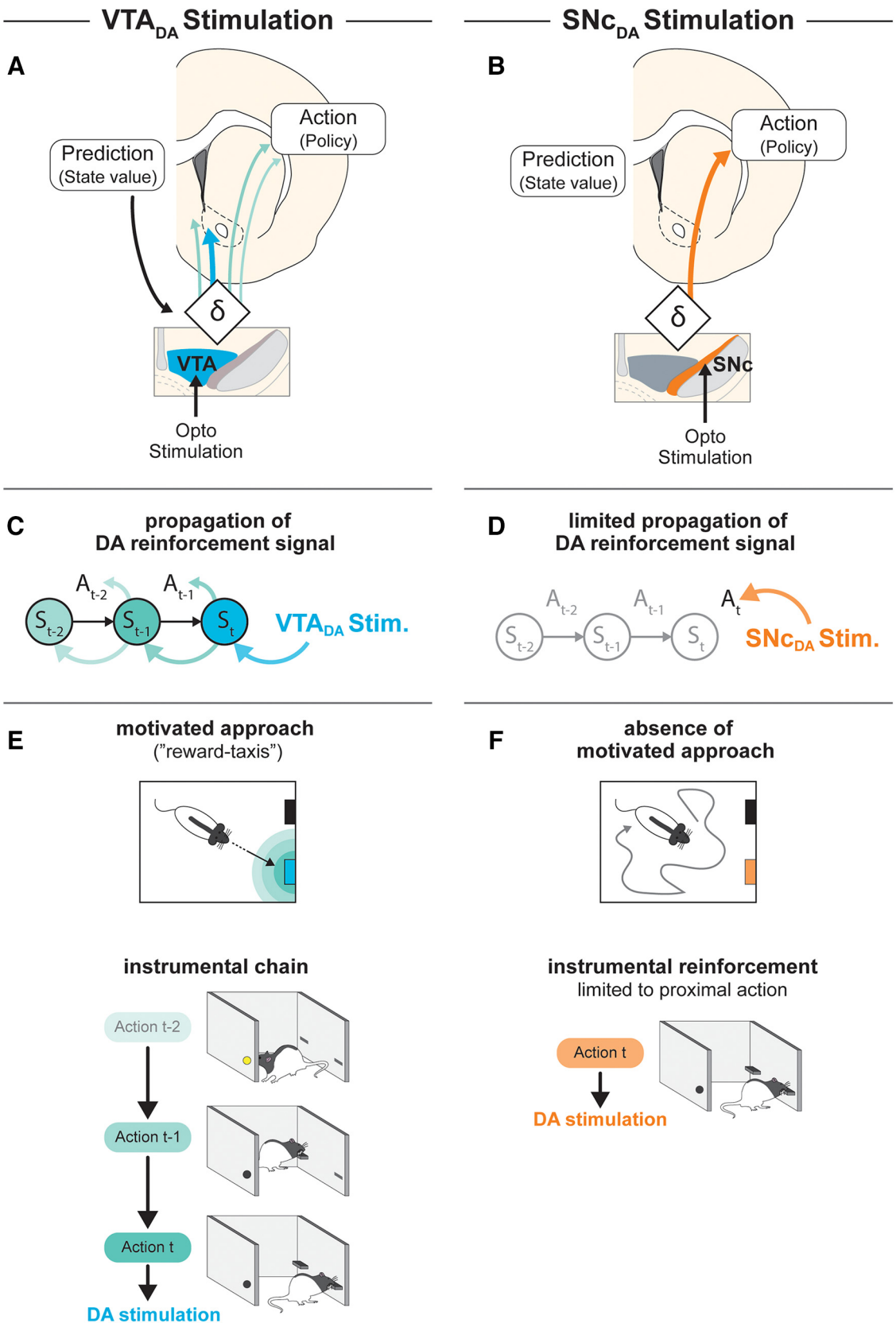

Figure 8. Proposed mechanism for VTA_{DA}-mediated and SNc_{DA}-mediated reinforcement. ***A***, ***B***, Actor–critic circuit motif. State values are proposed to be encoded in the ventromedial striatum (VMS; the critic). Policies or action values are proposed to be encoded in the dorsolateral striatum (DLS; the actor). VTA and SNc dopamine neurons provide reinforcement signals to VMS and DLS, respectively. Feedback projections from the ventromedial striatum allows valued states to activate VTA and SNc dopamine neurons resulting in the propagation of the dopamine reinforcement signal. For simplicity, the midbrain microcircuitry is not shown. ***C***, ***D***, This circuit motif predicts that stimulation of VTA_{DA} produces a direct reinforcement signal to the critic (blue arrow), but also temporally and anatomically propagated reinforcement signals to both the actor and the critic modules (green arrows). In contrast, stimulation of the SNc_{DA} produces a direct reinforcement signal to the actor (orange arrow) with little potential for backpropagation of that reinforcement signal. ***E***, ***F***, Consequences for behavior. See Discussion for details. Stim., Stimulation.

contingent sensory stimulation is known to maintain a non-negligible level of instrumental responding, even in absence of any other primary reward (Olsen and Winder, 2009; López et al., 2021). In and of themselves, these factors (type of operandum, novelty of the operandum, and ancillary cues) cannot explain the high rate of responding in VTA_{DA} or SNc_{DA} ICSS experiments, which is clearly because of the reinforcing properties of DA stimulation. However, by ensuring sporadic spontaneous interactions with the instrumental operandum, these experimental conditions might have compensated for the reduced motivation to initiate responding in SNc_{DA} rats and masked potential differences between VTA_{DA} and SNc_{DA} self-stimulation.

The timing of dopamine neuron activity and dopamine release is critical for neural plasticity and learning (Yagishita et al., 2014). Primary rewards induce burst activity in dopamine neurons and the release of dopamine for durations that range from a few milliseconds for small and discrete rewards, to several seconds for larger and sustained rewards (minutes in the case of drugs of abuse; Saunders et al., 2018; Ferguson et al., 2020; van Elzelingen et al., 2022). Likewise, the experimental activation of dopamine neurons has traditionally ranged from a few milliseconds (Coddington et al., 2023; Markowitz et al., 2023) to several seconds (Pascoli et al., 2015; Sharpe et al., 2017; Saunders et al., 2018; Hollon et al., 2021). Here we opted to stimulate dopamine neurons for 2 s (5 ms pulses, 20 pulses/s). These stimulation parameters approximate the endogenous activation of dopamine neurons and the striatal dopamine release observed in response to unexpected food pellets or a small bolus of sugar water, which are standard rewards in freely moving preparations (Saunders et al., 2018; van Elzelingen et al., 2022). Whether our results generalize to different stimulation parameters remains to be determined.

In this study, we interrogated VTA_{DA} and SNc_{DA} separately. However, these neural populations do not necessarily function independently from each other. Indeed, VTA_{DA} neurons can potentially influence SNc_{DA} neurons via ascending striato-nigro-striatal loops (Haber et al., 2000; Joel and Weiner, 2000; Wouterlood et al., 2018). Therefore, in this study the direct, targeted, optogenetic stimulation of VTA_{DA} neurons might have resulted in an indirect, propagated activation of the SNc_{DA} because of the spiraling interconnectivity between these regions. The extent to which the instrumental behavior observed during VTA_{DA} self-stimulation relies strictly on VTA_{DA} activation or reflects the additional recruitment of SNc_{DA} neurons remains to be determined. The potential for the learning-mediated temporal backpropagation of the dopamine reinforcement signal to cues and actions preceding stimulation following VTA_{DA}, but not SNc_{DA}, activation might contribute to the observed behavioral difference between VTA_{DA} and SNc_{DA} self-stimulation. Indeed, stimulation of VTA_{DA} and the resulting backpropagation of the dopamine reinforcement signal might allow for the temporally and spatially organized reward-seeking behavior observed during VTA_{DA} self-stimulation. In contrast, in the absence of a backpropagated dopamine signal, the reinforcing effect of SNc_{DA} would be limited to the elemental action that immediately precedes the stimulation (Hollon et al., 2021; Fig. 8). Rats self-stimulating SNc_{DA} neurons might therefore find themselves in a most peculiar situation, avidly engaging in instrumental responding when they, by chance, find themselves in proximity of the active lever, but otherwise showing little motivation to approach the lever or engage in the task.

In conclusion, consistent with prior studies we show here that the activation of either VTA or SNc dopamine neurons is a potent reinforcer of instrumental behavior (Rossi et al., 2013; Ilango et al., 2014; Saunders et al., 2018; Keiflin et al., 2019). Critically, however, we demonstrate that VTA_{DA} and SNc_{DA} neuron activation produce different "dimensions" of reinforcement, as reflected by the profound behavioral differences observed during VTA_{DA} and SNc_{DA} self-stimulation. Indeed, rats self-stimulating VTA_{DA} neurons demonstrated a flexible, organized, and motivated reward-seeking (i.e., stimulation-seeking) behavior. In contrast, rats self-stimulating SNc_{DA} neurons, while capable of a high rate of stereotyped instrumental behavior during response bouts (Hollon et al., 2021), appear to lack flexibility and motivation to initiate responding. Whether these behavioral differences (flexible vs rigid behavior) relate to habitual or goal-directed processes requires further exploration. Collectively, these results support the notion that the functional specialization of VTA and SNc dopamine neurons is not limited to Pavlovian learning but extends to the instrumental domain. Finally, these results highlight how these parallel, yet interacting, dopamine pathways might contribute to different levels of integration of operant behavior, from hierarchically organized action plans to elemental motor commands (Teitelbaum, 1977; Mogenson et al., 1980; Cooper and Shallice, 2000; Grafton and Hamilton, 2007; Keramati and Gutkin, 2013).

References

Barter JW, Li S, Lu D, Bartholomew RA, Rossi MA, Shoemaker CT, Salas-Meza D, Gaidis E, Yin HH (2015) Beyond reward prediction errors: the role of dopamine in movement kinematics. Front Integr Neurosci 9:39.

Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47:129–141.

Björklund A, Dunnett SB (2007) Dopamine neuron systems in the brain: an update. Trends Neurosci 30:194–202.

Bova A, Gaidica M, Hurst A, Iwai Y, Hunter J, Leventhal DK (2020) Precisely timed dopamine signals establish distinct kinematic representations of skilled movements. Elife 9:e61591.

Chang CY, Esber GR, Marrero-Garcia Y, Yau H-J, Bonci A, Schoenbaum G (2016) Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. Nat Neurosci 19:111–116.

Clemens KJ, Caillé S, Cador M (2010) The effects of response operandum and prior food training on intravenous nicotine self-administration in rats. Psychopharmacology (Berl) 211:43–54.

Coddington LT, Dudman JT (2018) The timing of action determines reward prediction signals in identified midbrain dopamine neurons. Nat Neurosci 21:1563–1573.

Coddington LT, Lindo SE, Dudman JT (2023) Mesolimbic dopamine adapts the rate of learning from action. Nature 614:294–302.

Cone JJ, Fortin SM, McHenry JA, Stuber GD, McCutcheon JE, Roitman MF (2016) Physiological state gates acquisition and expression of mesolimbic reward prediction signals. Proc Natl Acad Sci U S A 113:1943–1948.

Cooper R, Shallice T (2000) Contention scheduling and the control of routine activities. Cogn Neuropsychol 17:297–338.

da Silva JA, Tecuapetla F, Paixão V, Costa RM (2018) Dopamine neuron activity before action initiation gates and invigorates future movements. Nature 554:244–248.

Day JJ, Roitman MF, Wightman RM, Carelli RM (2007) Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat Neurosci 10:1020–1028.

de Jong JW, Afjei SA, Pollak Dorocic I, Peck JR, Liu C, Kim CK, Tian L, Deisseroth K, Lammel S (2019) A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. Neuron 101:133–151.e7.

Dodson PD, Dreyer JK, Jennings KA, Syed ECJ, Wade-Martins R, Cragg SJ, Bolam JP, Magill PJ (2016) Representation of spontaneous movement by dopaminergic neurons is cell-type selective and disrupted in parkinsonism. Proc Natl Acad Sci U S A 113:E2180–E2188.

Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, Koay SA, Thiberge SY, Daw ND, Tank DW, Witten IB (2019) Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature 570:509–513.

Ennaceur A, Delacour J (1988) A new one-trial test for neurobiological studies of memory in rats. 1: behavioral data. Behav Brain Res 31:47–59.
Enquist M, Lind J, Ghirlanda S (2016) The power of associative learning and the ontogeny of optimal behaviour. R Soc Open Sci 3:160734.
Ferguson LM, Ahrens AM, Longyear LG, Aldridge JW (2020) Neurons of the ventral tegmental area encode individual differences in motivational "wanting" for reward cues. J Neurosci 40:8951–8963.
Glimcher PW (2011) Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc Natl Acad Sci U S A 108 [Suppl 3]:15647–15654.
Grafton ST, Hamilton AFdC (2007) Evidence for a distributed hierarchy of action representation in the brain. Hum Mov Sci 26:590–616.
Grilly DM (1977) Rate-dependent effects of amphetamine resulting from behavioral competition. Biobehav Rev 1:87–93.
Grove JCR, Gray LA, La Santa Medina N, Sivakumar N, Ahn JS, Corpuz TV, Berke JD, Kreitzer AC, Knight ZA (2022) Dopamine subsystems that track internal states. Nature 608:374–380.
Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J Neurosci 20:2369–2382.
Hamilos AE, Spedicato G, Hong Y, Sun F, Li Y, Assad JA (2021) Slowly evolving dopaminergic activity modulates the moment-to-moment probability of reward-related self-timed movements. Elife 10:e62583.
Hollon NG, Williams EW, Howard CD, Li H, Traut TI, Jin X (2021) Nigrostriatal dopamine signals sequence-specific action-outcome prediction errors. Curr Biol 31:5350–5363.e5.
Howe MW, Dombeck DA (2016) Rapid signalling in distinct dopaminergic axons during locomotion and reward. Nature 535:505–510.
Ikemoto S (2007) Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. Brain Res Rev 56:27–78.
Ilango A, Kesner AJ, Keller KL, Stuber GD, Bonci A, Ikemoto S (2014) Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. J Neurosci 34:817–822.
Joel D, Weiner I (2000) The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. Neuroscience 96:451–474.
Karin O, Alon U (2022) The dopamine circuit as a reward-taxis navigation system. PLoS Comput Biol 18:e1010340.
Keiflin R, Pribut HJ, Shah NB, Janak PH (2019) Ventral tegmental dopamine neurons participate in reward identity predictions. Curr Biol 29:93–103.e3.
Keramati M, Gutkin B (2013) Imbalanced decision hierarchy in addicts emerging from drug-hijacked dopamine spiraling circuit. PLoS One 8:e61489.
Lerner TN, Shilyansky C, Davidson TJ, Evans KE, Beier KT, Zalocusky KA, Crow AK, Malenka RC, Luo L, Tomer R, Deisseroth K (2015) Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. Cell 162:635–647.
López AJ, Johnson AR, Kunnath AJ, Morris AD, Zachry JE, Thibeault KC, Kutlu MG, Siciliano CA, Calipari ES (2021) An optimized procedure for robust volitional cocaine intake in mice. Exp Clin Psychopharmacol 29:319–333.
Markowitz JE, Gillis WF, Jay M, Wood J, Harris RW, Cieszkowski R, Scott R, Brann D, Koveal D, Kula T, Weinreb C, Osman MAM, Pinto SR, Uchida N, Linderman SW, Sabatini BL, Datta SR (2023) Spontaneous behaviour is structured by reinforcement without explicit reward. Nature 614:108–117.
Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459:837–841.
Mekarski JE (1988) Main effects of current and pimozide on prepared and learned self-stimulation behaviors are on performance not reward. Pharmacol Biochem Behav 31:845–853.
Menegas W, Akiti K, Amo R, Uchida N, Watabe-Uchida M (2018) Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. Nat Neurosci 21:1421–1430.
Mirenowicz J, Schultz W (1996) Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature 379:449–451.
Mogenson GJ, Jones DL, Yim CY (1980) From motivation to action: functional interface between the limbic system and the motor system. Prog Neurobiol 14:69–97.
Nicola SM (2010) The flexible approach hypothesis: unification of effort and cue-responding hypotheses for the role of nucleus accumbens dopamine in the activation of reward-seeking behavior. J Neurosci 30:16585–16600.
Olds J, Milner P (1954) Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. J Comp Physiol Psychol 47:419–427.
Olsen CM, Winder DG (2009) Operant sensation seeking engages similar neural substrates to operant drug seeking in C57 mice. Neuropsychopharmacology 34:1685–1694.
Pascoli V, Terrier J, Hiver A, Lüscher C (2015) Sufficiency of mesolimbic dopamine neuron stimulation for the progression to addiction. Neuron 88:1054–1066.
Ramayya AG, Misra A, Baltuch GH, Kahana MJ (2014) Microstimulation of the human substantia nigra alters reinforcement learning. J Neurosci 34:6887–6895.
Richardson NR, Roberts DC (1996) Progressive ratio schedules in drug self-administration studies in rats: a method to evaluate reinforcing efficacy. J Neurosci Methods 66:1–11.
Roitman MF, Wheeler RA, Wightman RM, Carelli RM (2008) Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. Nat Neurosci 11:1376–1377.
Rossi MA, Sukharnikova T, Hayrapetyan VY, Yang L, Yin HH (2013) Operant self-stimulation of dopamine neurons in the substantia nigra. PLoS One 8:e65799.
Saunders BT, Richard JM, Margolis EB, Janak PH (2018) Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat Neurosci 21:1072–1083.
Schindler CW, Thorndike EB, Goldberg SR (1993) Acquisition of a nose-poke response in rats as an operant. Bull Psychon Soc 31:291–294.
Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.
Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599.
Shahan TA (2010) Conditioned reinforcement and response strength. J Exp Anal Behav 93:269–289.
Sharpe MJ, Chang CY, Liu MA, Batchelor HM, Mueller LE, Jones JL, Niv Y, Schoenbaum G (2017) Dopamine transients are sufficient and necessary for acquisition of model-based associations. Nat Neurosci 20:735–742.
Stauffer WR, Lak A, Yang A, Borel M, Paulsen O, Boyden ES, Schultz W (2016) Dopamine neuron-specific optogenetic stimulation in rhesus macaques. Cell 166:1564–1571.e6.
Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH (2013) A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci 16:966–973.
Stujenske JM, Spellman T, Gordon JA (2015) Modeling the spatiotemporal dynamics of light and heat propagation for in vivo optogenetics. Cell Rep 12:525–534.
Teitelbaum P (1977) Levels of integration of the operant. In: Handbook of operant behavior (Honig WK, Staddon JER, eds), pp 7–27. Englewood Cliffs, NJ: Prentice-Hall.
van Elzelingen W, Goedhoop J, Warnaar P, Denys D, Arbab T, Willuhn I (2022) A unidirectional but not uniform striatal landscape of dopamine signaling for motivational stimuli. Proc Natl Acad Sci U S A 119:e2117270119.
Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. Nature 412:43–48.
Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C, Stuber GD, Tye KM, Janak PH, Deisseroth K (2011) Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. Neuron 72:721–733.
Wouterlood FG, Engel A, Daal M, Houwen G, Meinderts A, Jordà Siquier T, Beliën JAM, van Dongen YC, Scheel-Krüger J, Thierry A-M, Groenewegen HJ, Deniau J-M (2018) Mesencephalic dopamine neurons interfacing the shell of nucleus accumbens and the dorsolateral striatum in the rat. J Neurosci Res 96:1518–1542.
Yagishita S, Hayashi-Takagi A, Ellis-Davies GCR, Urakubo H, Ishii S, Kasai H (2014) A critical time window for dopamine actions on the structural plasticity of dendritic spines. Science 345:1616–1620.
Yuan L, Dou Y-N, Sun Y-G (2019) Topography of reward and aversion encoding in the mesolimbic dopaminergic system. J Neurosci 39:6472–6481.