



Tonic or Phasic Stimulation of Dopaminergic Projections to Prefrontal Cortex Causes Mice to Maintain or Deviate from Previously Learned Behavioral Strategies

Ian T. Ellwood,^{1,2,3,4*} Tosha Patel,^{1,2,3,4*} Varun Wadia,^{1,2,3,4*} Anthony T. Lee,^{1,2,3,4} Alayna T. Liptak,^{3,5}  Kevin J. Bender,^{3,5} and  Vikaas S. Sohal^{1,2,3,4}

¹Department of Psychiatry, ²Weill Institute for Neurosciences, ³Kavli Institute for Fundamental Neuroscience, ⁴Sloan–Swartz Center for Theoretical Neurobiology, and ⁵Department of Neurology, University of California, San Francisco, California 94143

Dopamine neurons in the ventral tegmental area (VTA) encode reward prediction errors and can drive reinforcement learning through their projections to striatum, but much less is known about their projections to prefrontal cortex (PFC). Here, we studied these projections and observed phasic VTA–PFC fiber photometry signals after the delivery of rewards. Next, we studied how optogenetic stimulation of these projections affects behavior using conditioned place preference and a task in which mice learn associations between cues and food rewards and then use those associations to make choices. Neither phasic nor tonic stimulation of dopaminergic VTA–PFC projections elicited place preference. Furthermore, substituting phasic VTA–PFC stimulation for food rewards was not sufficient to reinforce new cue–reward associations nor maintain previously learned ones. However, the same patterns of stimulation that failed to reinforce place preference or cue–reward associations were able to modify behavior in other ways. First, continuous tonic stimulation maintained previously learned cue–reward associations even after they ceased being valid. Second, delivering phasic stimulation either continuously or after choices not previously associated with reward induced mice to make choices that deviated from previously learned associations. In summary, despite the fact that dopaminergic VTA–PFC projections exhibit phasic increases in activity that are time locked to the delivery of rewards, phasic activation of these projections does not necessarily reinforce specific actions. Rather, dopaminergic VTA–PFC activity can control whether mice maintain or deviate from previously learned cue–reward associations.

Key words: behavioral flexibility; dopamine; learning and memory; perseveration; prefrontal cortex

Significance Statement

Dopaminergic inputs from ventral tegmental area (VTA) to striatum encode reward prediction errors and reinforce specific actions; however, it is currently unknown whether dopaminergic inputs to prefrontal cortex (PFC) play similar or distinct roles. Here, we used bulk Ca^{2+} imaging to show that unexpected rewards or reward-predicting cues elicit phasic increases in the activity of dopaminergic VTA–PFC fibers. However, in multiple behavioral paradigms, we failed to observe reinforcing effects after stimulation of these fibers. In these same experiments, we did find that tonic or phasic patterns of stimulation caused mice to maintain or deviate from previously learned cue–reward associations, respectively. Therefore, although they may exhibit similar patterns of activity, dopaminergic inputs to striatum and PFC can elicit divergent behavioral effects.

Introduction

The prefrontal cortex (PFC) plays a particularly important role in behavioral flexibility, the ability to shift rapidly from a previously

learned action–reward association to a new one when the rules of a task change (Miller and Cohen, 2001). Earlier studies suggested

Received May 3, 2017; revised June 20, 2017; accepted July 14, 2017.

Author contributions: I.T.E. and V.S.S. designed research; I.T.E., T.P., V.W., A.T. Lee, A.T. Liptak, and K.J.B. performed research; I.T.E. analyzed data; I.T.E. and V.S.S. wrote the paper.

This work was supported by the Staglin Family, the International Mental Health Research Organization, the National Institute of Mental Health–National Institutes of Health (NIH) (Grant K01MH097841 to I.T.E. and Grants ROOMH085946 and R01MH100292 to V.S.S.), the National Institute on Drug Abuse–NIH (Grant R01DA035913 to K.J.B.), the National Institute of General Medical Sciences–NIH (Medical Scientist Training Grant GM07618 to A.T. Lee), and the Office of the Director–NIH (Grant DP2MH100011 to V.S.S.). We thank A. Doupe and M. Warden for

comments on the manuscript; A. Brumback, K. Cho, S. Gee, C. Hansen, F. Luongo, T. Martin, K. Meda, and M. Warden for helpful discussions; J. Vazquez-DeRose for help with microdialysis and HPLC; M. Warden for help with implantation surgeries; and L. Grosenick for help with photometry.

The authors declare no competing financial interests.

*I.T.E., T.P., and V.W. contributed equally to this work.

Correspondence should be addressed to Dr. Vikaas Sohal, Department of Psychiatry, University of California, 675 Nelson Rising Lane, San Francisco, CA 94143-0444. E-mail: vikaas.sohal@ucsf.edu.

DOI:10.1523/JNEUROSCI.1221-17.2017

Copyright © 2017 the authors 0270-6474/17/378315-15\$15.00/0

that prefrontal dopamine plays a central role in this process, enhancing either flexibility or perseveration depending on levels of dopamine and/or which types of dopamine receptors are activated (Seamans et al., 1998; Durstewitz et al., 2000; Seamans and Yang, 2004; Floresco et al., 2006; Stefani and Moghaddam, 2006; Durstewitz and Seamans, 2008; St Onge et al., 2011; Puig and Miller, 2012, 2015). Together with the observation that VTA dopamine neurons fire in two modes (Grace and Bunney, 1983, 1984; Grace, 1991; Overton and Clark, 1997; Lapish et al., 2007; Schultz, 2007), this led to the idea that a single set of dopaminergic fibers, originating in the ventral tegmental area (VTA), can exert opposing effects on PFC-dependent behavioral flexibility. Tonic firing (~5 Hz) is generally believed to occur under baseline conditions in the absence of unexpected events, whereas phasic bursts of spikes at frequencies >20 Hz may signal unexpected rewarding (Schultz et al., 1997; Bayer and Glimcher, 2005; Schultz, 2006; Cohen et al., 2012) and/or aversive events (Mantz et al., 1989; Gao et al., 1990; Coizet et al., 2006; Brischoux et al., 2009; Bromberg-Martin et al., 2010). It has been conjectured that these two modes of firing may lead to different levels of prefrontal dopamine that modulate PFC-dependent behavior differentially (Grace, 1991; Seamans and Yang, 2004). Specifically, according to the “dual-state theory of prefrontal dopamine function,” tonic firing releases moderate or lower levels of dopamine, which stabilizes a single pattern of behavior (Durstewitz and Seamans, 2008). This has also been referred to as the “exploit” mode of PFC function (Daw et al., 2006). Conversely, by releasing higher levels of dopamine, phasic firing is hypothesized to destabilize previously learned behavioral strategies, shifting the PFC to a more flexible mode of behavior.

The dual-state theory contrasts with the prevailing (non-PFC-specific) model of dopamine signaling, which suggests that phasic bursts in dopamine fibers elicit positive reinforcement (Berridge, 2007; Tsai et al., 2009; Adamantidis et al., 2011; Flagel et al., 2011; Schultz, 2013). For example, phasic bursts of activity in dopaminergic projections to the nucleus accumbens (NAc) reinforce associations between recently experienced cues and reward (Adamantidis et al., 2011; Steinberg et al., 2013). If VTA-to-PFC projections function similarly, then it is natural to assume that phasic bursts transmitted by these projections should also reinforce recent actions associated with unexpected rewards. Note, however, that one study suggests that dopaminergic VTA-to-PFC projections may transmit aversive signals (Lammel et al., 2012).

The difference between the dual-state and positive reinforcement models reflects, in part, the different experimental observations on which they are based. The dual-state model is based on pharmacological manipulations within the PFC, whereas the positive reinforcement model is motivated by the observation that midbrain dopamine neurons tend to signal reward prediction errors. However, there are very few experiments that have stimulated PFC-projecting dopamine fibers specifically and directly to determine what kinds of behavioral effects they can elicit.

Here, we explored this topic, first by stimulating dopaminergic VTA-to-PFC projections continuously using tonic or phasic patterns, and then by delivering single bursts of phasic stimulation time-locked to specific choices. We also use *in vivo* microdialysis and slice electrophysiology to compare dopamine levels and glutamatergic excitation elicited by tonic versus phasic patterns of activity, as well as fiber photometry to identify specific behavioral contingencies that drive phasic increases in VTA-to-PFC input.

Materials and Methods

Procedures. All experiments were conducted in accordance with procedures established by the Administrative Panels on Laboratory Animal Care at the University of California–San Francisco.

Injection of mice for Channelrhodopsin-2 (ChR2) or enhanced yellow fluorescent protein (eYFP) expression and implantation of fibers. All mice were C57BL/6 TH::Cre (line F112, www.gensat.org). Only male mice were used in the odor/texture discrimination and conditioned place preference (CPP) tasks, whereas a mixture of male and female mice were used for the slice, photometry, and microdialysis experiments. Cre-dependent expression was driven using a previously described adeno-associated virus (AAV) containing the fusion protein DIO-ChR2-eYFP, DIO-eYFP, DIO-GCaMP6s, or DIO-eGFP under the EF1 or synapsin promoter (Sohal and Huguenard, 2003; Atasoy et al., 2008; Tsai et al., 2009). We injected 1.0–1.5 μ l of $4\text{--}10 \times 10^{12}$ vg/ml virus into right VTA or bilaterally using methods described previously (Gee et al., 2012). Coordinates relative to bregma in millimeters were -2.58 AP, ± 0.5 ML, and -4.4 DV for all behavioral and slice experiments. For microdialysis and photometry, two injections of 750 nl were made at (-2.58 AP, 0.5 ML, -4.4 DV) and (-3.08 AP, 0.5 ML, -4.4 DV). At least 6 weeks were allowed for expression time in slice and photometry experiments and at least 8 weeks in microdialysis and behavioral experiments. For light stimulation, 200 μ m optical fibers (Doric Lenses) were implanted over medial PFC (mPFC) (1.7 AP, ± 0.35 ML, -2.25 DV), right NAc (0.75 ML, 1.3 AP, -4.0 DV), or right VTA (0.4 ML, -3.0 AP, -3.9 DV), whereas, for imaging, a 400 μ m optical fiber was implanted at 1.7 AP, 0.3 ML, -2.6 DV. Mice used for staining in the VTA were injected bilaterally in the VTA as described above with DIO-GFP and then, 8 weeks later, injected in the mPFC with 300 μ l of red retrobeads (Lumafuor) (1.7 AP, 0.3 ML, -2.75 DV) and perfused 5 d later.

Photometry. Our methods followed the techniques described in Gunaydin et al. (2014), 473 nm light was generated by a LuxX 473 nm laser diode (Market Tech), passed through an optical chopper (Thor Laboratories) running at 400 Hz and through a 473/10 nm laser clean-up filter (Semrock), reflected off of a 495 nm single-edge dichroic beam splitter (Semrock) and collimated into a 400 μ m fiber attached to the mouse's fiber-optic implant. Emitted light passed from the fiber through the beamsplitter and a 525/50 nm bandpass filter (Semrock) and was focused onto a femtowatt silicon photoreceiver (Newport). Signals from the photoreceiver were passed into a lock-in amplifier (Stanford Research Systems), which also received signals from the optical chopper (to determine the amplified frequency). The time constant was set at 3 ms.

In three animals, we also used a second 405 nm laser reflected off of a 427 nm beam splitter and passed through the same fiber to the mouse. In these experiments, instead of using an optical chopper, we used two pulse generators to modulate both the 473 and 405 nm lasers sinusoidally at 400 and 565.685 Hz, respectively. The output from the detector was then passed to two separate lock-in amplifiers to isolate the two signals.

After collection, signals were fit with, $f(t) = A + B \cdot \exp(C \cdot t)$, which was used to estimate $\Delta F/F$ at each time point. The resulting baselined signal was then band-pass filtered between 0.01 and 10 Hz using a 2-pole causal Butterworth filter. To determine the relative magnitude of the 405 and 473 nm laser signals, we performed a linear fit of the 405 laser signal versus 473 nm laser signal and used the resulting parameters to rescale and shift the 405 laser signal. Because we found a slow drift in the fit parameters over the 2 h of recording time, we performed this fit in 2 min windows around each time point.

Photometry task. The mice were trained to associate an LED and 500 ms 10 kHz tone with a water reward delivered from a lickometer. Licking before the reward cue triggered 500 ms of white noise and an overhead light. Training was performed in four phases. Mice were required to achieve at least 20 correct trials to proceed to the next phase of training and were returned to the previous phase if they received <10 correct trials. Phase 1 included a 30 s intertrial interval. There was no delay between cue and water reward and no punishment for licking between cues for 60 min total time. Phase 2 included a 30 s intertrial interval, a 3 s delay between cue and water reward, for 90 min total time. Phase 3 included 30–60 s between trials with a 6 s delay. Licking between trials

activated white noise (500 ms) and an overhead “house light” (10 s) as well as resetting the intertrial interval for 2 h total time. Phase 4 was the same as experimentation day, with a 60–120 s inter trial interval, a 10 s delay, and the same effect of licking between trials as phase 3 for 2 h total time. Before training, mice were day/night shifted and water deprived for 3 d until they reached 80% of baseline weight and were given at least 700 μ l of water per day to maintain their weight.

CPP. We tested for a CPP using a 6 d paradigm consisting of a habituation day and a pretest day in which mice were allowed to explore a 3-chamber custom CPP box freely, followed by 3 d of conditioning, during which mice received stimulation, and a test day. During the conditioning days, the center chamber of the CPP box was closed off from the outer chambers and mice received stimulation on one side and no light on the other side in a manner balanced across mice. Each session, habituation, pretest, test, and conditioning lasted 20 min. During conditioning days, mice were run in the stimulation and no light conditions in the morning or afternoon with at least 3 h between runs. Following a method described previously (Roux et al., 2003), mice that spent >75% of their time in the outer chambers on one side during the pretest were not included in the study. Mice bilaterally implanted in the mPFC were run first with phasic stimulation and then, in a separate chamber, with different patterns on the chamber wall and in a different room run with tonic stimulation.

Odor/texture discrimination task. All mice were housed in a day/night-reversed facility starting 4–6 d before experimentation. In addition, the experimenter was blinded as to whether the animals were expressing DIO-ChR2-eYFP or DIO-eYFP. Beginning at this time, the mice were food restricted and maintained at 80–85% of initial weight, whereas water was available *ad libitum*. Throughout the food restriction period (no more than 2 weeks), food was only available in 2 bowls at one end of the cage and consisted of 1–3 g of chopped Reese’s peanut butter chips (Hershey) buried under the same digging medium used in the task. The digging medium consisted of one texture, either white sand (Mosser Lee) or bicarbonate-free cat litter (Cole Valley Pets), and one odor (1% by volume), either ground coriander seed (McCormick) or garlic powder (McCormick). In addition, 1.3 mg/ml of finely chopped peanut butter chip was added to the medium.

After a mouse reached its target weight (3–4 d), it was placed in a holding cage while two bowls were prepared in its home cage from one of the two digging medium groups with a food reward in one of the bowls (~10 mg peanut butter chip). The mouse was then returned to its home cage and allowed to explore both bowls freely until it had both found the reward and explored the unrewarded bowl. This was repeated at least eight and up to 20 times until the mouse found the reward rapidly on at least four consecutive trials. During these trials, a fiber-optic cable was attached to the mouse’s implant, but no light stimulation was delivered.

Mice then received 1 or 2 d of training in the main task as described in the Results section. During training, the mice were always trained on one rule and then switched to another rule, whereas during experiment days, a number of variations on the task were used, as described in the Results. Mice were attached to a fiber-optic cable, but received no stimulation.

A single task trial proceeded as follows. The mouse was placed in a holding cage while two bowls were filled with medium and one was baited with a reward. The rewards were always placed according to the current “rule,” which was either a fixed texture or odor associated with the reward location. Whether the reward was on the left or right and which group of digging medium was used on a given trial was determined randomly ahead of time with repetitions of the same direction or group longer than three removed.

The mouse was only allowed to dig in one of the two bowls (either with its feet or its nose) and was removed to the holding cage if it: (1) received the reward, (2) gave up exploring the bowl that had no reward, or (3) did not dig in either bowl for 2 min. After a successful trial in which the mouse received the reward, a new trial was begun (~1 min per trial). If the mouse selected the bowl without reward, the bowl with reward was removed and the mouse was allowed to explore the unrewarded bowl until it lost interest. The mouse was then placed in the holding cage for an additional 1 min (~2 min per trial). If the mouse showed no interest in either bowl, the mouse was removed to the holding cage and the next trial

was started after a 1 min delay and the trial was scored as a “time-out.” A mouse was deemed to have learned a rule if it received the reward on eight of the 10 previous trials. Mice unable to learn the initial and second rule after 2 d of training were removed from the study.

Stimulation protocols. For behavior and microdialysis experiments, light stimulation was delivered via fiber-optic cable fed through a commutator (Doric Lenses) and attached to a 100 mW, 473 nm laser (OEM) driven by a pulse generator. For unilaterally implanted animals, the total light power delivered during a pulse was 5 or 15 mW for mice injected with 4×10^{12} and 1×10^{13} titer virus. Bilaterally implanted animals were injected with 7.4×10^{12} titer virus and 3–5 mW was delivered on each side for the bowl-digging task and CPP. For microdialysis, as the mice were anesthetized and to deliver light to all sides of the probe, 20 mW was used. For slice experiments, 3 or 6 mW of 470 nm light was delivered via a DG4 xenon arc lamp (Sutter Instruments) through a 40 \times objective on an Olympus BX51WI microscope.

Slice experiments. Slice preparation and intracellular recordings followed a previously established protocol (Sohal and Huguenard, 2005). Slices were cut to 250 μ m from 10- to 11-week-old male and female mice and bathed ACSF containing the following (in mM): 126 NaCl, 26 NaHCO₃, 2.5 KCl, 1.25 NaH₂PO₄, 1 MgCl₂, 2 CaCl₂, and 10 glucose. Whole-cell recordings were obtained using an internal solution containing the following (in mM): 130 K-gluconate, 10 KCl, 10 HEPES, 10 EGTA, 2 MgCl₂, 2 MgATP, and 0.3 NaGTP pH adjusted to 7.3 with KOH. Slices were secured using a harp, with the harp strings placed so as to avoid mPFC. Neurons in layer VI of prelimbic and infralimbic cortex were identified visually using differential contrast video microscopy on an upright microscope (BX51WI; Olympus). Recordings were performed at $32 \pm 1^\circ\text{C}$ using a Multiclamp 700A (Molecular Devices) and patch electrodes with resistance 2–4 M Ω . Series resistance was typically 10–20 M Ω and recordings were discarded >30 M Ω . The glutamate receptor antagonists CNQX and APV (Tocris Bioscience) were bath applied and delivered through the perfusion system.

Microdialysis. Anesthetized mice (1% isoflurane at 0.6 L/min) were acutely implanted with a combined microdialysis (CMA) and light-fiber probe (Doric Lenses) in mPFC. The dialysis probe had a 2 mm, 5000 kD cutoff membrane and was perfused throughout the experiment at 1 μ l/min. After implantation, the probe was perfused for 2 h before samples were collected. Six 20 μ l samples were collected and analyzed per animal: a baseline sample, a tonic/phasic stimulation sample, another baseline sample, and, after a 20 min delay, another baseline sample, phasic/tonic stimulation sample, and baseline sample. Tonic and phasic stimulated samples were compared with the average of their neighboring baseline samples. The perfusion fluid consisted of the following (in mM): 148.1 NaCl, 3 KCl, 1.4 CaCl₂ · 2H₂O, 0.8 MgCl₂ · 6H₂O, 0.8 Na₂HPO₄ · 7H₂O, and 0.2 NaH₂PO₄ · H₂O. Samples were collected in vials containing 5 μ l of 0.3 M perchloric acid, which were kept on ice. After collection, samples were immediately frozen at -80°C until being shipped on dry ice to SRI International for analysis. One mouse was not included in the study because we were not able to maintain a constant level of 1% isoflurane during the procedure. A second mouse in which the amount of perchloric acid was doubled in an attempt to preserve more dopamine was excluded due to a 30% drop-off in measured dopamine levels in the baseline samples during the experiment, suggesting that the probe may have been compromised.

Histology and immunohistochemistry. Mice were anesthetized with pentobarbital or avertin tribromoethanol and perfused with 4% PFA in ice-cold PBS. After removal, brains were fixed overnight in 4% PFA before being transferred to a 30% sucrose solution. Then, 40–50 μ m slices were obtained using a Leica VT 1200S vibratome or cryostat. The primary antibodies used were chicken anti-TH (1:1000, Millipore 9702), 1 $^\circ$ /2 $^\circ$ conjugate anti-GFP (Alexa Fluor 489), sheep anti-TH (1:1000, Abcam 113), rabbit anti-GABA (1:1000, Abcam 9446), and chicken anti-GFP (1:1000, Aves Labs 1020). The secondary antibodies used were goat anti-chicken (1:500, Alexa Fluor 569), donkey anti-sheep (1:500, Alexa Fluor 546), goat anti-rabbit (1:500, Alexa Fluor 647), and goat anti-chicken (1:500, Alexa Fluor 488). Slices were mounted in 0.3% gelatin and placed under a coverslip with Vectashield and DAPI and imaged using a confocal microscope (Leica 510) or an upright microscope

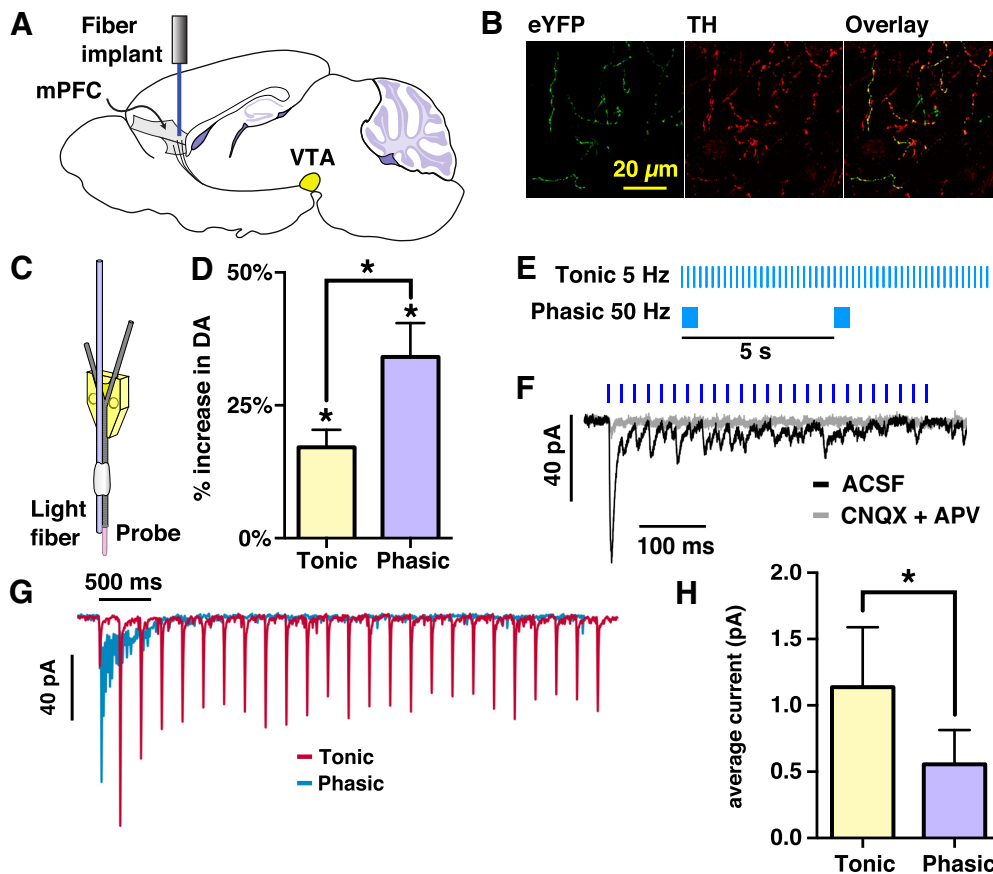


Figure 1. Effects of stimulating dopaminergic fibers from the VTA to PFC and mouse performance in rule-changing task. $*p < 0.05$. **A**, AAV5-EF1-DIO-ChR2-eYFP or AAV5-EF1-DIO-eYFP was injected into right VTA. For mice used in behavioral experiments, a fiber-optic cannula was also implanted above mPFC. **B**, Slice of mPFC showing eYFP, TH stain, and overlay. For additional histology, see also Figure 1-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f1-1>). **C**, Combined fiber and microdialysis probe used to collect dialysis samples. **D**, Percentage increase (relative to baseline) of extracellular dopamine under the two stimulation conditions. Both tonic and phasic stimulation showed a significant increase ($n = 4$, $p = 0.014$ tonic, $p = 0.013$ phasic, Student's t test, $t = 5.24$ tonic, $t = 5.34$ phasic), with phasic releasing more dopamine than tonic ($p = 0.024$, $n = 4$, Student's t test, $t = 4.27$). **E**, Illustration of the continuous tonic and phasic patterns of stimulation. Two periods of stimulation are shown. **F**, Example voltage-clamp recording from a layer VI neuron during phasic light stimulation of dopaminergic fibers (average of 5 runs). Responses in ACSF are shown in black. Responses after 15 min of bath applied CNQX ($5 \mu\text{M}$) and APV ($25 \mu\text{M}$) are shown in gray. **G**, Example recording of excitatory currents during phasic (blue) and tonic (red) stimulation. **H**, Comparison of inward current during tonic and phasic stimulation averaged over the stimulation period. Tonic stimulation induced significantly more current than phasic ($p = 0.0312$, $n = 7$ neurons, Wilcoxon signed-rank test, $W = 1$).

(Nikon Eclipse 80i). Coexpression was determined using four $230 \times 230 \mu\text{m}$ images each of the VTA and PFC taken at $40\times$. Due to the punctate nature of the TH stain, only PFC fiber segments longer than $5 \mu\text{m}$ were examined. Furthermore, to ensure accurate counting of mPFC-projecting neurons in the VTA, only cells with four or more beads in the cell body were included in the analyses.

Statistics and sample size. All error bars shown represent the SEM unless otherwise stated in the figure legends. All statistical tests performed were two-sided. Mann–Whitney U tests; Wilcoxon signed-rank tests were used in comparisons of the number of correct trials when the number of trials was <30 for any of the datapoints being tested because the distributions were discrete and non-Gaussian. When all runs had 30 trials, Student's t -test used. A Wilcoxon signed-rank test was also used when comparing the amount of current induced in layer VI cells by light stimulation of TH fibers because many of the currents were small, with a few large outliers making their distribution non-Gaussian.

The sample size for the continuous tonic and phasic stimulation experiments was set ahead of time to be at least 10 ChR2 animals and at least four eYFP animals, which were run in batches of two to five animals at a time with one to two eYFP animals per cohort except the first cohort of two ChR2 animals. Experiments in which the food reward was omitted were added to the tests that each animal was given after the effects of stimulation on a rule shift and maintenance of the initial association were found. We implanted nine ChR2 and nine eYFP mice bilaterally, of which eight and seven, respectively, were usable for experiments in which

we delivered stimulation on incorrect trials without a rule shift and stimulation on correct trials without food. These mice were run in five cohorts of three to four mice with varying numbers of eYFP and ChR2 mice. One eYFP mouse in this cohort was only used for stimulation on incorrect because its implant was damaged between runs. Eight mice were implanted for the bilateral single phasic burst during rule shift experiments and were run in three cohorts (two or three ChR2 mice and one eYFP mouse). The VTA-implanted mice were run in a single cohort of four ChR2-injected mice.

Results

We injected heterozygous TH-Cre mice with virus to drive Cre-dependent expression of either ChR2-eYFP or eYFP in the VTA (Atasoy et al., 2008; Sohal et al., 2009; Tsai et al., 2009) and implanted them with an optical fiber (Doric Lenses) over the right medial mPFC (Fig. 1A and Fig. 1-1A, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f1-1>). We waited at least 8 weeks for expression before behavioral experiments. Initially, our injection and implantations were both unilateral on the right side. Subsequently, we also performed bilateral injections and implantations, as described below.

First, we confirmed that this viral and transgenic approach mainly labels dopaminergic neurons. In the PFC, 88.4% of

eYFP⁺ fibers also costained for TH (129/146 fibers). To further characterize the labeled VTA neurons that project to mPFC, we also injected a small number of heterozygous TH-Cre mice with virus to drive Cre-dependent GFP expression in the VTA, along with retrogradely transported fluorescent microspheres (Retrobeads) in the mPFC. Similar to the result of staining eYFP⁺ projections in mPFC for TH, we found that, within the VTA, 88.6% of GFP⁺ neurons that were labeled with retrobeads also costained for TH (140/158 neurons; Fig. 1-1B, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f1-1>). Among the 143 neurons that contained retrobeads and were TH⁺, 140 were GFP⁺. We also examined the small fraction (~11%) of GFP-labeled, mPFC-projecting VTA cells that were TH⁻. We found that 50% of these cells costained for GABA (9/18 neurons). In summary, in our TH-Cre mice, virtually all (98%) of mPFC-projecting, TH⁺ VTA neurons at the injection site were labeled successfully. Conversely, we estimate that, among the mPFC-projecting VTA neurons we were stimulating, ~90% stained for TH, whereas 5% lacked TH but stained for GABA. We note that, whereas the majority of cells stimulated in this study were TH⁺ neurons, it is possible that some of effects that we describe may be modified by the remaining 10% of TH⁻ neurons, which were activated and recorded.

Stimulating VTA–mPFC projections releases dopamine and glutamate in the mPFC

Next, we confirmed that optogenetic stimulation of labeled VTA–mPFC fibers releases dopamine. We initially explored the effects of two patterns of light stimulation (Tsai et al., 2009) modeled after the native firing of dopamine neurons (Fig. 1E). For “tonic” stimulation we used a steady 5 Hz train of 4 ms pulses, whereas for “phasic” stimulation, we used a 50 Hz burst of 4 ms pulses lasting 500 ms and occurring every 5 s. Therefore, both patterns of stimulation delivered a total of 25 light pulses every 5 s. Notably, this phasic frequency has been found to be close to the ideal frequency for maximum dopamine release in the striatum, ~40–50 Hz (Bass et al., 2010).

To measure dopamine release elicited by optogenetic stimulation of these VTA projections, we used a combination optical fiber (Doric Lenses) and microdialysis probe (CMA) in anesthetized mice (Fig. 1C). Microdialysis samples were collected over 20 min periods of stimulation with the average of the 20 min before and after stimulation used as a baseline. Samples were then analyzed using HPLC (SRI). We found significant increases in dopamine levels with both phasic stimulation (34.1% increase, $p = 0.013$, $n = 4$, Student's t test, $t = 5.34$) and tonic stimulation (17.1% increase, $p = 0.014$, $n = 4$, Student's t test, $t = 5.24$). Phasic stimulation released significantly more dopamine than tonic stimulation ($p = 0.024$, $n = 4$, Student's t test, $t = 4.27$) (Fig. 1D).

Finally, before beginning behavioral experiments, we also used slice electrophysiology to determine whether optogenetic stimulation of these VTA projections could elicit synaptic responses in mPFC neurons, as suggested by the fact that many PFC-projecting dopamine neurons costain for VGluT2 (Gorelova et al., 2012). In layer VI neurons held in voltage clamp at -70 mV, we found stimulus-locked excitatory responses in 16/46 cells. These EPSCs were reduced significantly after bath applying CNQX (5–10 μ M) and APV (25–50 μ M) for 15 min ($n = 4$ cells, $p = 0.018$ Student's t test, $t = -4.74$) (Fig. 1F). Notably, among the 16 neurons in which we found responses to light stimulation, 14 had responses within 2–6 ms after the onset of light stimulation (mean 3.9 ± 0.3 ms), suggestive of direct responses. Examining only neurons with light-evoked responses larger than 10

pA, the average inward current during tonic stimulation was significantly larger than during phasic stimulation, suggesting that clustering the light flashes into phasic bursts reduced the overall level of glutamatergic excitation ($p = 0.031$, $n = 7$ neurons, Wilcoxon signed-rank test, $W = 1$) (Fig. 1G,H).

GCaMP recordings from VTA–mPFC dopaminergic fibers

Having confirmed that VTA–mPFC projections release dopamine, we next sought to identify behavioral conditions under which these projections are activated. In particular, generic VTA dopamine neurons typically exhibit phasic bursts when an animal receives an unexpected reward or unexpected cue that predicts a reward and we wanted to determine whether mPFC-projecting dopamine neurons exhibit a similar response profile. To investigate this, TH-Cre mice were injected with virus to drive Cre-dependent expression of GCaMP6s or (as a control) eGFP in the right VTA, and implanted over mPFC with a 400 μ m optical fiber (Doric Lenses) (Fig. 2-1A, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f2-1>). After waiting 5–6 weeks, mice were water deprived and trained in an operant chamber task in which a brief LED light and 10 kHz tone (500 ms) was followed after 10 s by a 10 μ l water reward delivered from a lickometer on the chamber wall (Fig. 2A,B). The intertrial interval was a randomly selected time between 1 and 2 min. To encourage mice to form an association with the tone/LED cue, touching the lickometer before the cue presentation triggered a mild “punishment” consisting of 500 ms of white noise, an overhead light (10 s), and a reset of the 1–2 min intertrial interval (timeout). After 5–7 d of training, mice were attached via an optical fiber to a photometry rig (Adelsberger et al., 2005; Cui et al., 2013, 2014; Gunaydin et al., 2014) and GCaMP fluorescence signals were measured over 2 h. The measured signals were small, but could be resolved by averaging over many trials (Fig. 2C; for single-trial examples, see Fig. 2-1B, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f2-1>).

As seen in Figure 2D, GCaMP6s signals from VTA–mPFC dopaminergic projections rose immediately after the tone/light cue that predicted reward. Notably, there was no cue-evoked rise in GFP-expressing control mice (Fig. 2E) (change in $\% \Delta F/F$ from 10 s before the cue to a window 0.5–1.5 s after the cue: 0.62 ± 0.14 for GCaMP6s vs 0.029 ± 0.012 for GFP; $n = 25$ mice with GCaMP and 8 with GFP; $p = 0.004$, Mann–Whitney U test, $U = 66$). After premature licks, we observed no significant change in fluorescence (change in $\% \Delta F/F = 0.038 \pm 0.025$ for GCaMP6s vs -0.004 ± 0.008 for GFP, $p = 0.30$, Mann–Whitney U test, $U = 111$). There was a second, much smaller, increase in fluorescence at the time of reward delivery (Fig. 2D; change in $\% \Delta F/F$ from 2 s before reward to a window 0.5–3.5 s after reward delivery: 0.094 ± 0.084 for GCaMP6s vs -0.034 ± 0.020 for GFP mice; $p = 0.03$, Mann–Whitney U test, $U = 84$).

In addition to comparing GCaMP fluorescence with signals from GFP-expressing animals, we performed a second control experiment to rule out possible artifactual sources of signal by delivering 405 and 473 nm light at two different temporal frequencies to stimulate GCaMP6s, then separating the responses to these two different excitation wavelengths using two lock-in amplifiers (Lerner et al., 2015). The 405 nm light causes GCaMP6s to fluoresce even in the absence of Ca^{2+} ; therefore, response to 405 nm excitation serves as a control for possible activity-independent changes in the photometry signal such as movement-related changes. As seen in Figure 2F and Figure 2-1B (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f2-1>), the cue-evoked signal was absent when measuring 405 nm-driven fluorescence,

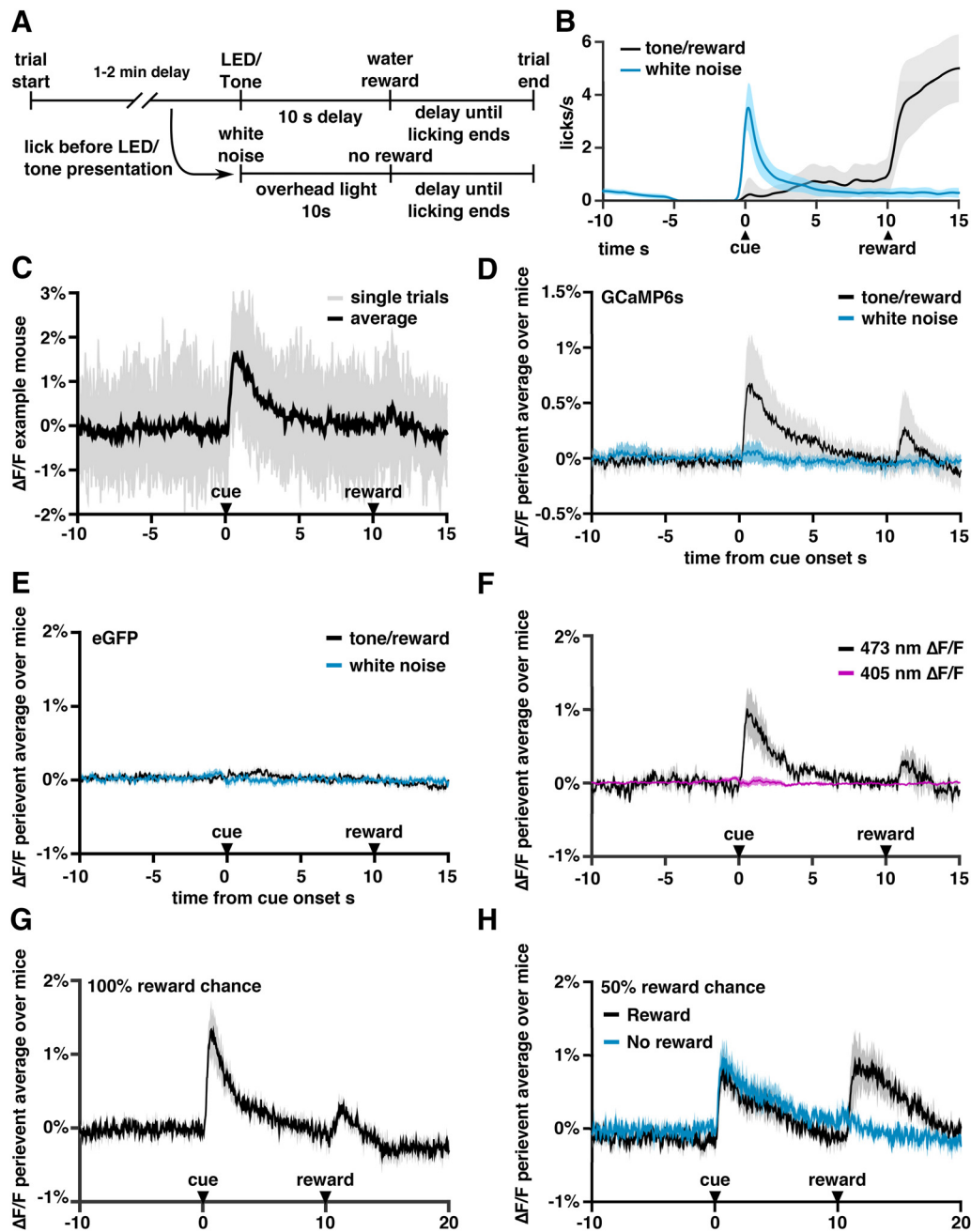


Figure 2. Recordings of activity in dopaminergic fibers in the PFC using fiber photometry. For additional data from the recordings, see also Figure 2-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f2-1>). **A**, Timeline of operant chamber task. **B**, Average licks/s on trials in which the mouse received the tone/LED cue (black) or prematurely licked triggering white noise and an overhead light (blue trace). Light gray and blue regions indicate SEM. Note that the peak in the blue trace at $t = 0$ is a consequence of the fact that there must be at least one lick at $t = 0$ to trigger the white noise/overhead light response. **C**, Recordings from an example mouse. Single trials are shown in gray and the trial average in black. Note the small signal-to-noise ratio. **D**, Average fluorescence across all DIO-GCaMP6s injected mice of perievent average $\Delta F/F$ around cue presentation. Black line represents trials in which the tone/light cue was presented and a water reward delivered 10 s later. Blue line represents trials in which the mouse licked without a tone presentation, triggering white noise, an overhead light, and resetting the intertrial delay. **E**, Average fluorescence across DIO-GFP injected mice. **F**, Average fluorescence across mice excited with both 473 nm light and 405 nm light. **G**, Average fluorescence during a modified task in which rewards were only delivered on 50% of trials after a cue. Black indicates trials in which the reward was delivered; blue, trials in which no reward was delivered.

confirming that the cue-evoked GCaMP signal was not due to artifacts such as movement.

Having found cue- and reward-evoked responses, we investigated whether we could measure VTA-mPFC signals related to reward omission. For this, we selected six mice that exhibited robust signals on the preceding paradigm ($>0.5\%$ deviations in $\Delta F/F$) and trained them on a new paradigm in which the water reward was only delivered after a tone/LED cue on 50% of trials.

Before training on this new paradigm, these mice showed cue and reward responses that were similar to the averages across our entire cohort (Fig. 2G). However, during the 50% reward condition, we found a significant increase in the reward-evoked responses compared with previous condition in which mice always received rewards (Fig. 2H and Fig. 2-1C, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f2-1>) (change in $\% \Delta F/F$ from its average 2 s before reward, to its average during the period

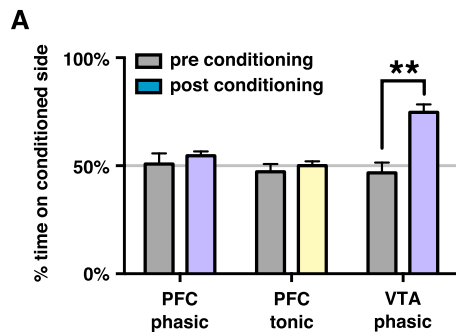


Figure 3. Phasic and tonic stimulation do not lead to conditioned place preference/aversion. **A**, Mice bilaterally injected with DIO-ChR2 in the VTA and implanted with a fiber optic over mPFC showed no significant place preference in a CPP paradigm using either continuous phasic or continuous tonic stimulation. Mice implanted with a fiber optic over the VTA were also tested with phasic stimulation and formed a significant preference for the conditioned side.

0.5–3.5 s after reward: 0.1 ± 0.07 for 100% reward ($n = 6$) vs 0.7 ± 0.2 for 50% reward task, $n = 6$, $p = 0.03$, Wilcoxon signed-rank test, $W = 21$). In contrast, at the time of reward omission, we observed a significant decrease in the GCaMP signal (change in $\% \Delta F/F = -0.1 \pm 0.02$, $p = 0.03$, Wilcoxon signed-rank test, $W = 2$). It is important to note that, even in this case, the signal still did not dip below the baseline level of fluorescence. In summary, we find changes in $\Delta F/F$ after cues that predict rewards and after the delivery of rewards. However, we did not observe increases or decreases in $\Delta F/F$ associated with the absence of an expected reward or cues indicating a delay until the next reward. These findings suggest that, at least in the absence of strongly aversive (e.g., painful) cues, increases in activity of mPFC-projecting dopamine fibers encode positive reward prediction errors. Although we did not find a dip in the signal when rewards were denied, this may have been due to the inability of our fiber photometry approach to resolve changes in fluorescence beneath an already small baseline signal.

Continuous tonic and phasic VTA–mPFC stimulation fails to elicit place preference/aversion

Having found that VTA–mPFC projections exhibit phasic reward-related activity, we decided to test whether stimulating these projections was rewarding. Specifically, we assayed the effects of continuous tonic or phasic stimulation using a CPP paradigm (Fig. 3A). As a positive control, we used mice implanted with an optical fiber over the VTA instead of the mPFC. Neither tonic nor phasic VTA–mPFC stimulation elicited significant preference or avoidance for the conditioned side (change in percentage time on conditioned side for phasic stimulation: $3.8 \pm 5.6\%$, $n = 7$ mice, $p = 0.52$, Student's t test, $t = 0.682$; for tonic stimulation: $-2.9 \pm 5.2\%$, $n = 8$ mice, $p = 0.60$, Student's t test, $t = 0.547$). In contrast, phasic stimulation within the VTA elicited significant preference for the conditioned side (change in percentage time on conditioned side = $27.8 \pm 4.3\%$, $n = 5$ mice, $p = 0.003$, Student's t test, $t = 6.53$). Popescu et al. (2016) also found that phasic VTA–mPFC stimulation is not sufficient to encourage licking behavior, consistent with this result. Importantly, these same patterns of continuous tonic or phasic VTA–mPFC stimulation, which failed to affect CPP, had marked effects on behavior within an mPFC-dependent behavioral flexibility task, described below.

Odor/texture discrimination task for measuring behavioral flexibility

To study how VTA–mPFC stimulation affects behavioral flexibility, we used a task developed previously (Cho et al., 2015), a simplified version of previous odor/texture discrimination tasks (Birrell and Brown, 2000; Bissonette et al., 2008; Bissonette and Powell, 2012) in which mice choose to dig in one of two bowls to find a food reward (Fig. 4A). Each bowl was marked by two cues: one digging medium (either sand or litter) and one odor (either coriander or garlic). On each trial, the two bowls had different odors and media (Fig. 4B). The food rewards were always associated with one odor or one digging medium. Therefore, there were four different food–stimulus associations, or “rules”, that the mouse could learn and a mouse that chose randomly had a 50% chance of getting a reward on each trial. Mice did not show a strong preference for selecting the left or right bowl in any of our experiments (Fig. 4-1A, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f4-1>) and we performed additional tests to ensure that they were unable to detect the reward directly without using the cues (Fig. 4-1B, C, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f4-1>). Because of the small number of odor and texture cues, this task should not be considered a true “set-shifting” task because mice learned specific rules, but did not necessarily learn to attend specifically to just one modality. However, this task is mPFC dependent (Cho et al., 2015) and the simplicity was advantageous because mice were typically able to perform the task on the first day of testing.

A mouse was considered to have learned a rule if it selected the correct bowl on eight of 10 trials and we kept the rule constant until the mouse either met this criterion or failed to do so after 30 trials. After a mouse reached this criterion for one rule, we performed three additional trials using the original rule before testing their ability to switch to a second rule, which, unless otherwise noted, was from the other set of cues (e.g., if the first rule associated reward with an odor, then the second rule would associate a texture with reward). This task design was critical because, after the rule change, errors could be classified as either “perseverative,” consistent with the initial rule, or “random,” inconsistent with both the initial and new rules (Fig. 4C). As described below, this made it possible to distinguish between different types of impairments in learning the new rule.

We tested TH-Cre mice that had been injected with virus to drive Cre-dependent expression of ChR2 in the VTA and implanted with an optical fiber over the mPFC. Mice readily learned the initial association (mean 13.2 ± 0.3 trials to criterion, 15 mice, $n = 95$ runs; Fig. 4D) and a new association after a rule change (mean 15.1 ± 0.8 trials to criterion, 15 mice, $n = 15$ runs; Fig. 4E). On the first four trials after the rule change, mice continued to make choices consistent with the initial association $\sim 80\%$ of the time (Fig. 4E, blue line). However, 10 trials after the rule change, mice selected bowls consistent with the old rule at chance levels.

We note that, in the subsequent discussion, we will typically plot only the percentage correct or percentage perseverative average across mice as a function of trial number; however, in these plots, when mice reach criteria, they drop out of our averages, which can create artifacts when the number of trials to criterion is highly variable. In such cases, where averaging over mice poorly represents the data, we will present only the unaveraged data, which can be found in the extended data figures for all of the experiments.

We performed numerous experiments, described below, using this task paradigm. First, we studied how the continuous delivery or tonic or phasic patterns of stimulation affected the

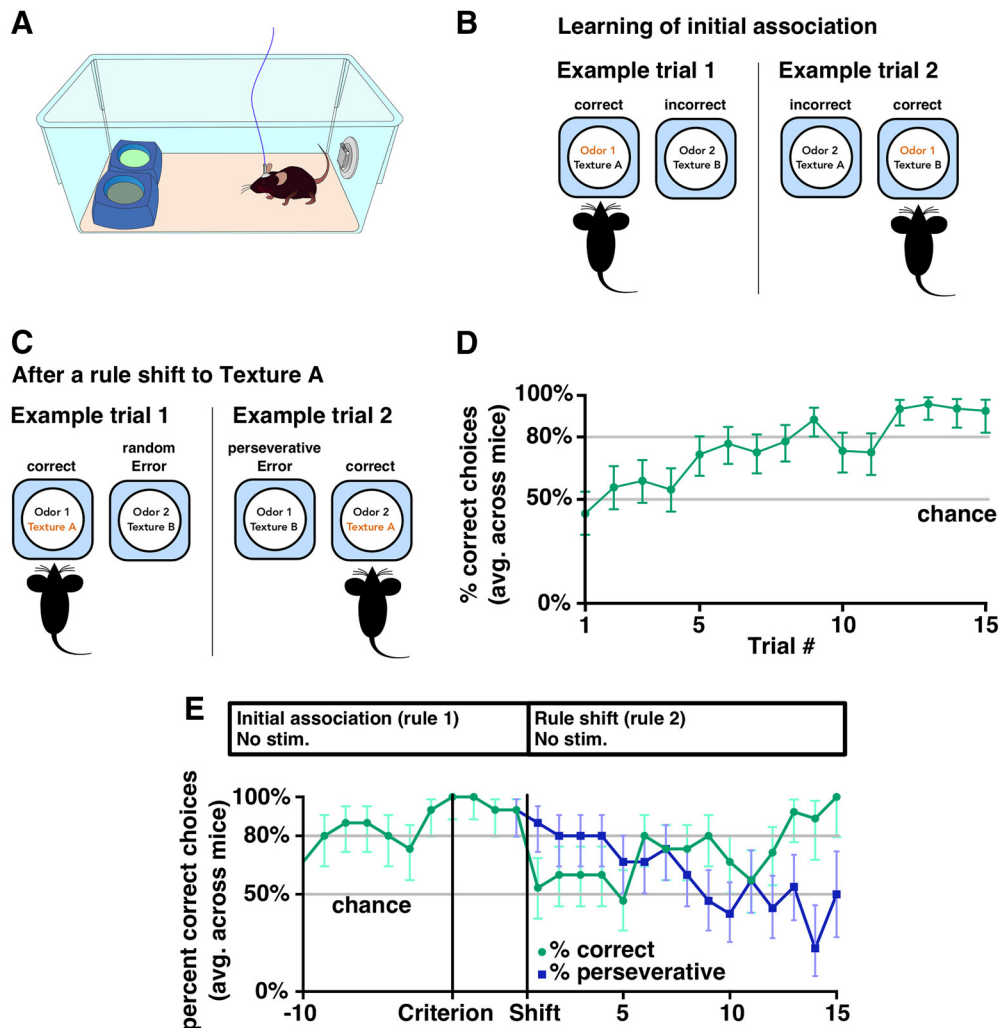


Figure 4. Odor/texture discrimination task for measuring behavioral flexibility. For additional details of the task, see also Figure 4-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f4-1>). **A**, Task configuration at the beginning of each trial. **B**, Two example trials in which one of the odors (odor 1) is selected as the rule. Note that odor 1 can be paired with either of the two textures. Selecting the correct bowl yielded a food reward. **C**, Example trials after a rule shift from one of the odors to one of the textures. Selecting a bowl with texture A yielded a food reward, whereas selecting the other bowl was classified as either a perseverative or random error depending on whether the choice was consistent with the previous rule (odor 1). **D**, Percentage of trials in which a mouse dug in the correct bowl during an initial association, averaged across 95 runs performed by 15 mice (includes all initial association periods from experiments with unilaterally implanted mice). 95% Clopper–Pearson confidence intervals are shown. **E**, Performance after a rule shift averaged across mice unilaterally injected with ChR2 or eYFP run without light stimulation. Blue line indicates the percentage of trials in which the mouse was perseverative (i.e., selected a bowl based on the first rule). See also Figure 4-1B (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f4-1>). Because we performed many variations of this task, we have included a task outline above each plot similar to this one. These diagrams show the stages of the experiment, along with various modifications to the original design such as whether the mice received light stimulation. The first rectangle represents the period in which mice learned an initial association and ends when they reach criterion. Any three-trial buffer period is not shown for simplicity. The second rectangle represents the subsequent trials in which, in this case, the food reward is delivered after rule 2 instead of rule 1.

ability of mice to switch to a new association after they already had learned an initial association. We studied the types of errors (perseverative vs random) that mice made in each case. Then, we studied the effects of tonic or phasic stimulation on the ability to maintain a previously learned association. Next, we studied the behavioral effects of single phasic bursts delivered after specific choices. We specifically investigated whether single phasic bursts are sufficient to reinforce specific actions and how single phasic bursts alter the ability of mice to maintain a previously learned association or to switch to a new one when delivered after correct or incorrect choices.

Continuous tonic and phasic stimulation disrupt learning of a new association (but do so in distinct ways)

As outlined above, we studied how tonic and phasic stimulation affect performance after a switch from an initial association to a

new one. Once animals reached the 80% criterion for the initial association, we began continuous tonic or phasic stimulation and continued testing on the initial association for three additional trials before switching to the new association. Both phasic and tonic stimulation had profoundly negative effects on learning of the new association (Fig. 5A–C and Fig. 5-1A–F, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f5-1>). All ChR2-expressing mice (11/11) receiving phasic stimulation and 9/10 ChR2-expressing mice receiving tonic stimulation were unable to reach the learning criterion within our cutoff of 30 trials (Fig. 5C), whereas all eYFP-expressing control mice were able to reach the learning criterion in the presence of either phasic (mean 14.2 ± 1.3 trials, 5 mice) or tonic (mean 14.8 ± 1.6 trials, 5 mice) stimulation.

Phasic and tonic stimulation both led to poor switching, but caused very different types of impairments. We defined the percent-

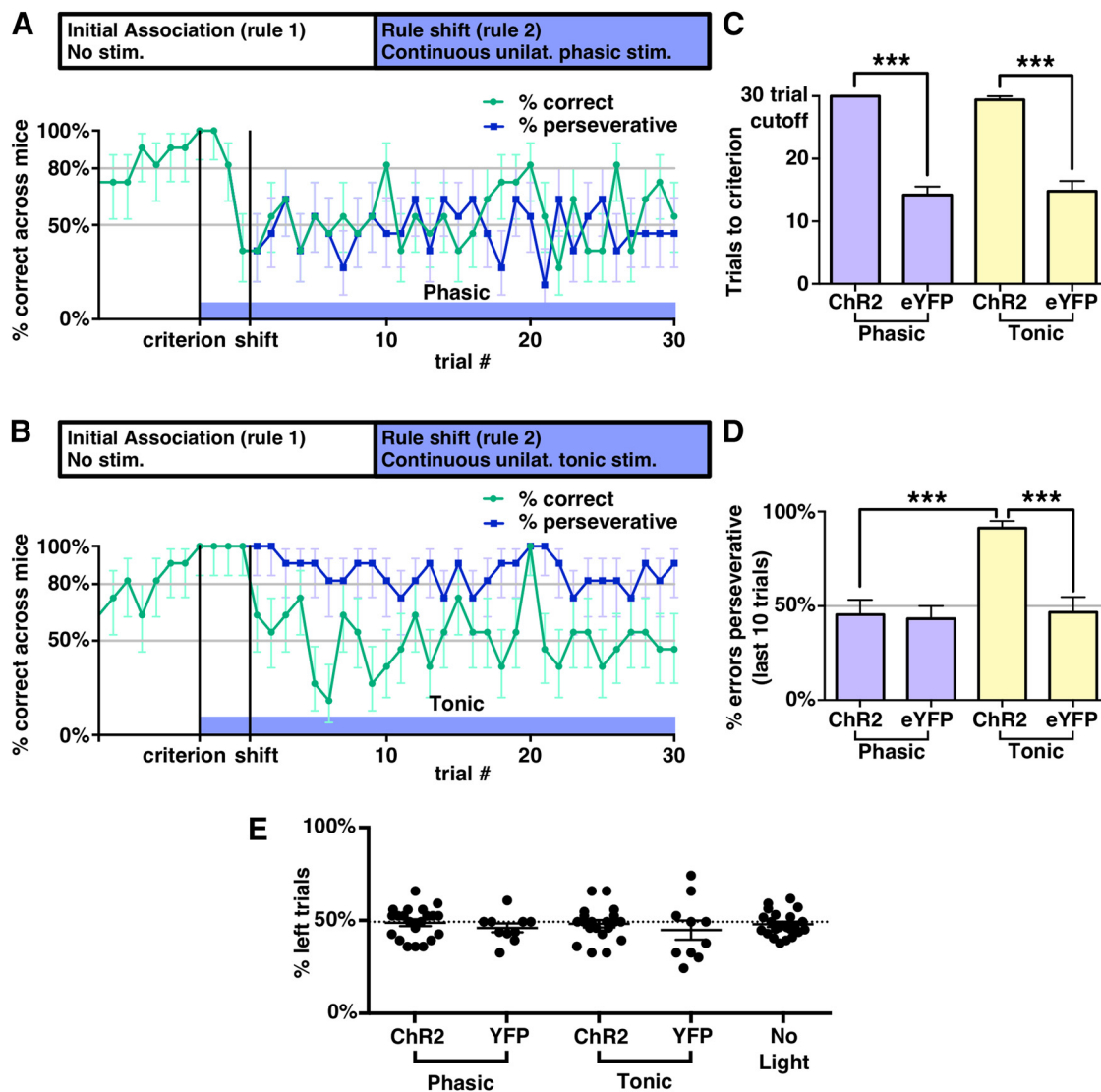


Figure 5. Tonic and phasic stimulation affect behavioral flexibility differentially after a rule shift. See also Figure 5-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f5-1>) for unaveraged data from individual mice and Figure 5-2 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f5-2>) for additional tests of the effects of tonic and phasic stimulation. **A**, Percentage of correct trials averaged across phasically stimulated mice. Blue bar indicates period of phasic stimulation, which began after mice reached criteria and continued until the end of the experiment. Blue line represents the percentage of perseverative choices. The second vertical line indicates when the rule was changed. Error bars were calculated using a Clopper–Pearson confidence interval with $\alpha = 31.8\%$, representing the SEM for binomially distributed variables. **B**, Same as **A**, but for tonic stimulation. **C**, Number of trials to criterion was significantly higher for ChR2-expressing mice compared with eYFP-expressing mice during both phasic stimulation ($p = 0.00046$, Mann–Whitney U test, $n = 11$ ChR2, 5 eYFP, $U = 121$) and tonic stimulation ($p = 0.00067$, Mann–Whitney U test, $n = 10$ ChR2, 5 eYFP, $U = 105$). **D**, Percentage of error trials in which the mouse made perseverative choices, averaged across the last 10 trials of the task and across mice. Tonic stimulation of ChR2-expressing mice were significantly more perseverative than expected by chance ($p = 0.0020$, Wilcoxon signed-rank test, $n = 10$, $W = 55$), or observed in either tonically stimulated eYFP-expressing mice ($p = 0.00067$, Mann–Whitney U test, $n = 10$ ChR2, 5 eYFP, $U = 105$) or phasically stimulated ChR2-expressing mice ($p = 0.00041$, Mann–Whitney U test, $n = 11$ phasic, 10 tonic, $U = 160$). **E**, Percentage of trials in which the left bowl was selected in the trials after three-trial buffer. Each dot represents a single run in that stimulation type. See also Figure 4-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f4-1>).

age of “perseverative” trials after the rule change as the percentage of choices that were consistent with the initial association (Fig. 5A,B,D). During the final 10 trials of the task, tonically stimulated ChR2-expressing mice were perseverative on $91 \pm 4\%$ of incorrect trials compared with $46 \pm 8\%$ for tonically stimulated eYFP-expressing mice ($p = 0.00067$, Mann–Whitney U test, $U = 105$). In contrast, phasically stimulated mice showed no preference for either the old or new association and made perseverative errors at chance levels ($45 \pm 7\%$ of the time). These effects appear to be specific for VTA projections to the mPFC because they were not reproduced when we used the same tonic and phasic patterns to stimulate TH⁺ fibers originating from the VTA in the NAc (Fig. 5-1G, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f5-1>).

We noted in Figure 5A that phasic stimulation affected performance before the rule shift and further experiments revealed that, even in the absence of a rule shift, phasically stimulated ChR2-expressing mice were unable to maintain an association during phasic stimulation, unlike eYFP-expressing mice (Fig. 5-2A, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f5-2>), (0/10 ChR2-expressing mice were able to re-reach criterion in 30 trials after stimulation onset vs average of 8.4 ± 0.2 trials to criterion for eYFP-expressing mice, $p = 0.00067$, Mann–Whitney U test, $n = 10$ ChR2, $n = 5$ eYFP, $U = 105$). We also observed that tonically stimulated mice would continue to prefer bowls based on the initial association even if all food rewards were omitted after the association had been learned (Fig.

5-2B, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f5-2>). Over 30 trials, tonically stimulated ChR2-expressing mice continued to select bowls based on the initial association $89 \pm 2\%$ of the time ($n = 5$ mice), whereas eYFP-expressing mice only made choices consistent with the initial association on $50 \pm 2\%$ of trials ($n = 3$ mice) ($p = 6.7 \times 10^{-6}$, Student's t test, $t = 14.5$). In addition, once we began omitting rewards, 3/3 eYFP-expressing mice had trials on which they failed to choose either bowl within 2 min, whereas this never occurred for ChR2-expressing mice ($p = 0.036$, Mann–Whitney U test, $U = 21$). Finally, we observed no directional bias induced by either tonic or phasic stimulation (Fig. 5E).

Single phasic bursts of VTA–mPFC stimulation are not sufficient to reinforce specific actions

Next, we used bilateral stimulation to explore the behavioral effects of single, precisely timed phasic bursts (as opposed to continuous phasic stimulation). In particular, our photometry recordings showed that VTA–mPFC dopaminergic fibers exhibit phasic signals after cues which predict rewards, smaller signals after the delivery of expected rewards, and larger signals after rewards that are not completely predictable. Numerous theoretical and experimental studies have shown how this kind of “reward prediction error” signal can be used to drive reinforcement learning. However, as described earlier, neither phasic nor tonic stimulation elicits CPP. Therefore, we now performed three additional experiments to test whether delivering single phasic bursts of stimulation to dopaminergic VTA–mPFC fibers could reinforce specific associations in our odor/texture discrimination task.

First, we attempted to train mice on an initial association, but omitted food rewards and instead delivered a single phasic burst whenever a mouse selected the “correct” bowl (Fig. 6A and Fig. 6-1A,B, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f6-1>). As a positive control, we also performed this experiment in mice injected with ChR2 and implanted with a fiber directly over the VTA. As seen in Figure 7, A and C, mice stimulated in the mPFC performed at chance levels and were not able to form an association. In contrast, all mice stimulated directly in the VTA were able to form an association within our 30 trial cutoff (17.8 ± 2.0 trials for VTA-implanted mice, $n = 4$, vs 30 trials for all mPFC-implanted mice, $n = 4$, $p = 0.03$, Mann–Whitney U test, $U = 26$).

Although phasic VTA–mPFC stimulation was not sufficient to reinforce learning of an initial association, we wondered whether phasic VTA–mPFC stimulation might be able to reinforce the learning of a new association in mice that had already learned an initial association. We thus trained mice using a food reward until they reached criterion on a given rule and then tested

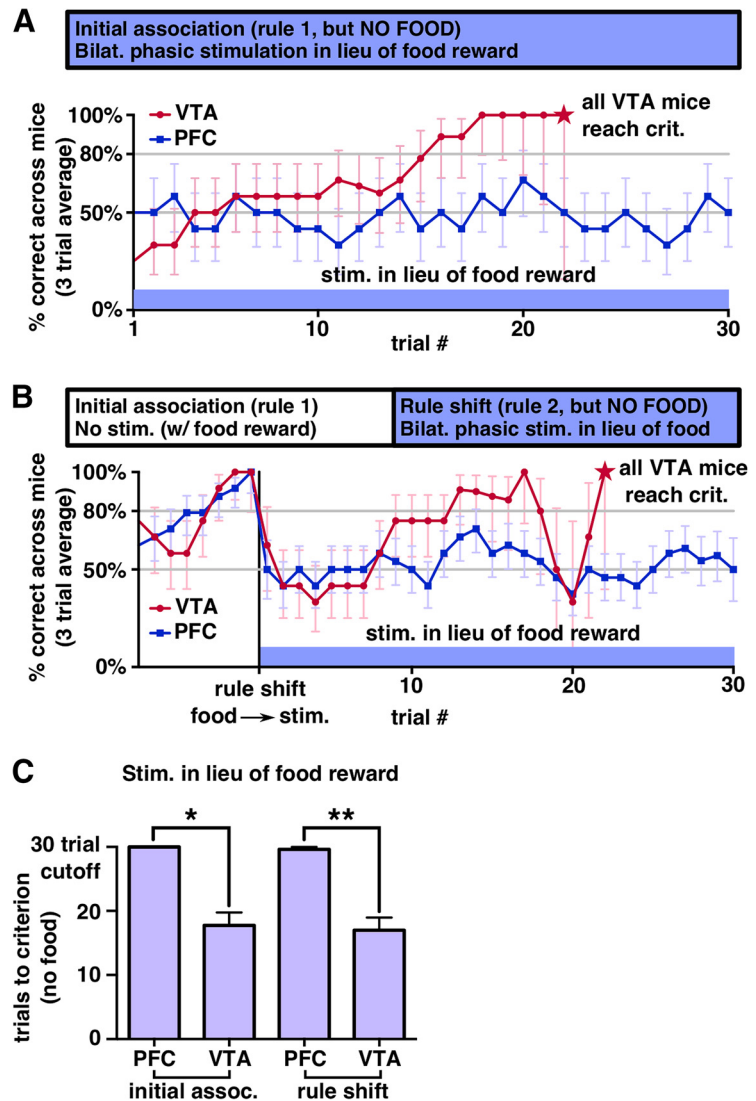


Figure 6. Phasic stimulation cannot be used in lieu of a food reward to learn new associations. For unaveraged data from individual mice, see also Figure 6-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f6-1>). **A**, Mice were tested on whether they would reach criteria in the odor/texture discrimination task if the food reward was replaced by a single phasic burst. Blue shows percentage correct for mice bilaterally implanted over mPFC, red shows mice implanted over VTA. **B**, Same as **A**, but mice were first trained in an initial association using a food reward and then tested to see if they would reach criterion in a new rule with light stimulation replacing the food reward. **C**, VTA implanted mice performed significantly better than PFC implanted mice in both the initial association and rule shift with light stimulation instead of food.

whether they could learn a new rule using single phasic bursts instead of a food reward. Here, we did find that one of the eight mice we ran was able to reach criterion in 27 trials, but we were unable to reproduce this effect in any of the other mice, suggesting that this one mouse may have succeeded by chance. Indeed, Figure 6, B and C (see also Fig. 6-1C,D, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f6-1>) shows that, on average, mice receiving bilateral VTA–mPFC phasic stimulation only after correct choices (with no food reward) guessed at chance levels, whereas VTA implanted mice were consistently able to learn the new association (17.0 ± 2.0 trials to criterion for VTA implanted mice, $n = 4$, vs 29.6 ± 0.4 trials for mPFC implanted mice, $n = 8$, $p = 0.004$, Mann–Whitney U test, $U = 10$).

Finally, we tested whether single bursts of phasic stimulation could substitute for food rewards to maintain a previously learned rule. For this, mice were again trained on an initial association. After reaching criterion, we then replaced the food reward with a single

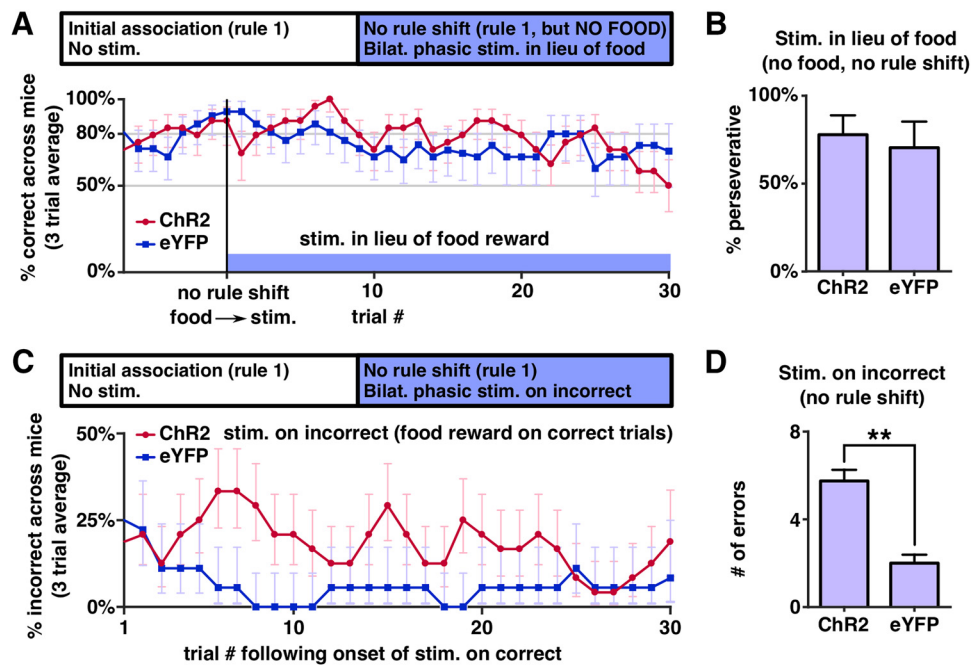


Figure 7. Phasic stimulation does not maintain associations when used in lieu of a food reward, but can disrupt existing associations. For unaveraged data from individual mice, see also Figure 7-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f7-1>). **A**, After meeting the criterion on an initial association formed using a food reward, mice were tested on whether they would maintain the association if the food reward was replaced by a single phasic burst. Red line shows Chr2 mice; blue line shows eYFP mice. **B**, There was no significant difference in the percentage of perseverative trials after the rule shift and replacement of the food reward with phasic stimulation between the Chr2 and eYFP injected cohorts. **C**, Mice were again trained on an initial association using a food reward. After meeting the criterion, the food reward continued to follow the initial association, but additional phasic bursts were delivered on incorrect trials. **D**, Stimulation on incorrect trials caused Chr2 injected mice to make significantly more errors over 30 trials than eYFP mice.

phasic burst delivered when the mouse selected the bowl that was “correct” based on the previously learned association. Again, trials were scored as a “timeout” when mice went >2 min without digging in either bowl. The experiment ended if mice made two consecutive timeouts. This pattern of VTA–mPFC phasic stimulation did not significantly improve maintenance of the previously learned rule because Chr2-injected mice and eYFP-injected mice made a similar fraction of perseverative choices ($n = 8$ Chr2 mice, 7 eYFP mice, $77.9 \pm 3.4\%$ for Chr2 mice vs $70.5 \pm 5.9\%$ for eYFP mice, $p = 0.44$, Mann–Whitney U test, $U = 71$) (Fig. 7A,B and Fig. 7-1A, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f7-1>).

Single phasic bursts of VTA–mPFC stimulation can increase choices that deviate from a previously learned rule

If phasic bursts of activity in dopaminergic VTA–mPFC projections do not serve to reinforce specific actions, then what is their function? Our previous findings suggest an intriguing possibility. Using fiber photometry, we observed that these phasic signals were strongest when rewards could not be predicted with complete certainty. Furthermore, continuous phasic stimulation elicits choices that are random with respect to side (left vs right), the previously learned rule, and the new rule being learning. Therefore, we hypothesized that phasic activity of VTA–mPFC dopaminergic projections might increase choices that deviate from the previously learned rule and that this effect might occur specifically when phasic VTA–mPFC activity follows actions that have not previously been associated with consistent rewards.

To test this possibility, we studied mice that had learned an initial association and delivered single phasic bursts specifically after errors; that is, choices that were not previously associated with reward. Specifically, TH-Cre mice injected with virus to drive Cre-dependent expression of either Chr2 or eYFP were

trained on an initial association and then, after reaching the learning criterion, tested on their ability to maintain that association over 30 trials while we delivered single phasic bursts whenever mice made an incorrect choice. Mice continued to receive food rewards after correct choices. We found that Chr2-expressing mice made almost 3 times the number of errors as control (eYFP-expressing) mice ($n = 8$ Chr2 mice, $n = 6$ eYFP mice, 5.8 ± 0.7 errors in Chr2 mice vs 2 ± 0.4 errors in eYFP mice, $p = 0.002$, Mann–Whitney U test, $U = 83$) (Fig. 7C,D and Fig. 7-1B, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f7-1>). Importantly, although the stimulation led to a significant increase in the number of errors, the chance of making errors remained low (~10–30%). Together with our previous findings, this suggests that single phasic bursts are not sufficient to strongly reinforce specific actions, but are sufficient to facilitate the sampling of choices that deviate from the previously learned association.

During the acquisition of new rules, appropriately timed phasic bursts do not disrupt learning

As described above, we found that phasic bursts of activity in VTA–mPFC dopaminergic projections are not sufficient to reinforce a new association. However, when these bursts occur after actions that were not previously associated with reward, they can increase choices that deviate from the previously learned association. This raises the following question: once an animal begins engaging in “exploration,” if rewards are associated with a new set of stimuli, are phasic bursts of VTA–mPFC activity compatible with learning of that new association or will they disrupt such learning? To address this question, we studied mice that had learned an initial association. We delivered phasic stimulation in two different ways while testing the ability of these mice to learn a new associa-

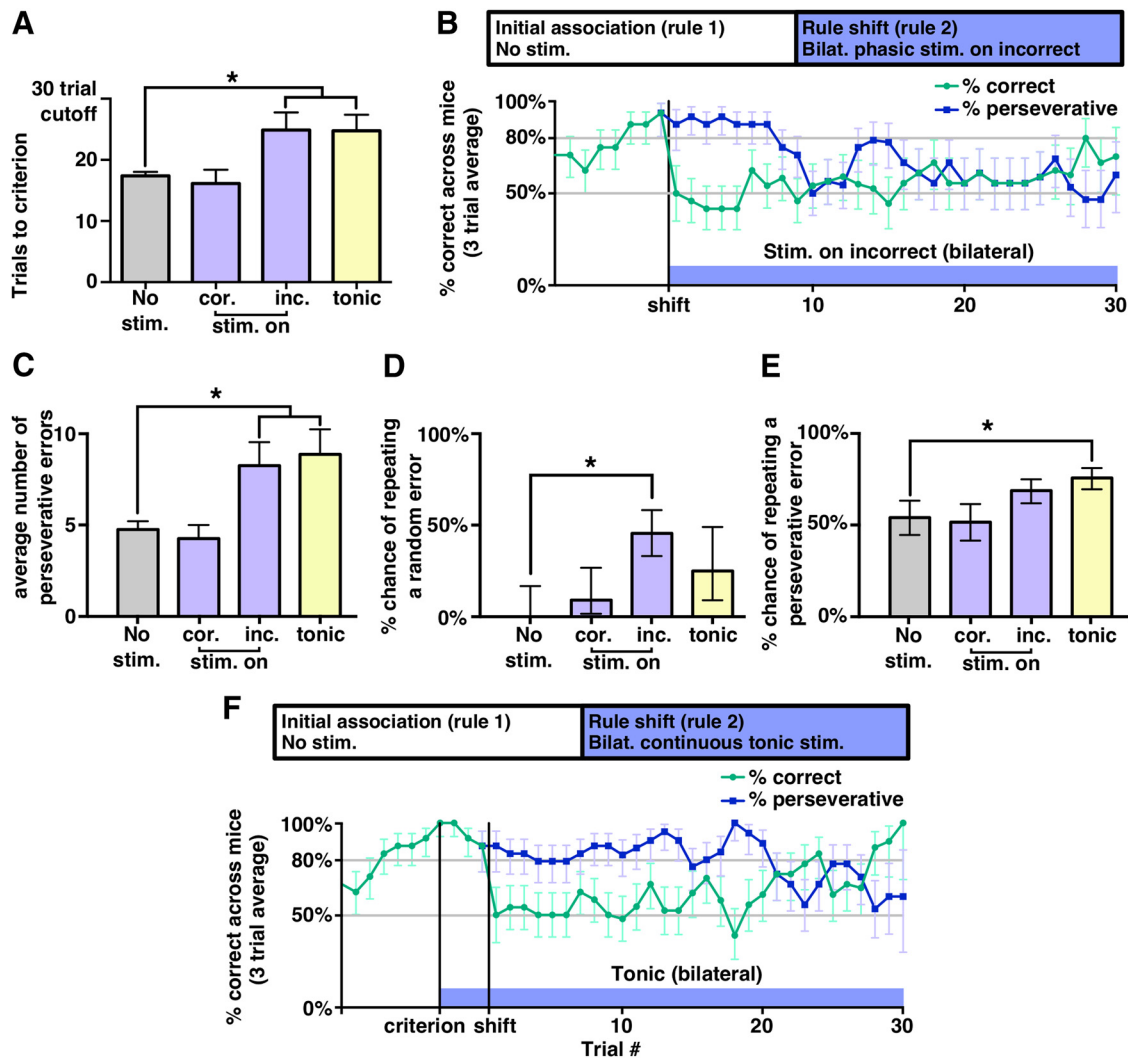


Figure 8. Bilateral phasic stimulation does not disrupt acquisition of a new rule when the stimulation is paired with the food reward, but does disrupt learning when paired with incorrect trials. Bilateral tonic stimulation produces perseveration. For unaveraged data from individual mice, see also Figure 8-1 (available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f8-1>). **A**, Mice required significantly more trials to reach criterion when stimulation was delivered on incorrect trials, but learned normally when stimulation was paired with the food reward. Tonic stimulation also significantly prolonged learning relative to mice receiving no light stimulation. **B**, Percentage correct averaged across mice for stimulation on incorrect. **C**, Both tonic stimulation and stimulation on incorrect showed significant increases in the number of perseverative trials. **D**, Only mice stimulated on incorrect trials showed a significantly increased chance of repeating a random error. **E**, Only mice stimulated with tonic stimulation showed a significantly increased chance of repeating perseverative errors. **F**, Percentage correct averaged across mice for bilateral tonic stimulation.

tion. In the first experiment, we delivered a single phasic burst of stimulation after each choice that was correct based on the new association. Mice continued to receive food after each correct choice. We observed that bilateral phasic stimulation delivered on correct trials was associated with learning in a similar number of trials compared with the control condition (no stim) ($n = 8$ mice, 17.4 ± 0.7 trials in control vs 16.1 ± 2.3 in stimulation on correct, $p = 0.74$, Wilcoxon signed-rank test, $W = 21$) (Fig. 8A and Fig. 8-1A–D, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f8-1>).

Phasic bursts after incorrect choices during a rule shift elicit disorganized behavior

Next, we again studied mice that had learned an initial association and began delivering stimulation as they shifted to a new rule. In this case, we delivered a single, bilateral phasic burst after each incorrect choice during the rule change. Again, correct choices were associated with food rewards. Stimulating on incorrect trials during learning of a new association led to a significant

increase in the number of trials needed to reach the learning criterion (Fig. 8A, B and Fig. 8-1E, F, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f8-1>) ($n = 8$ mice, 17.4 ± 0.7 trials in control vs 24.9 ± 2.9 in stimulating on incorrect, $p = 0.03$, Wilcoxon signed-rank test, $W = 3$).

The pattern of errors in mice receiving bilateral phasic stimulation after incorrect trials during a rule change was complex. There was an increase in perseverative errors ($n = 8$ mice, 4.8 ± 0.5 perseverative errors in control, 8.3 ± 1.3 for stim on incorrect, $p = 0.03$, Wilcoxon signed-rank test, $W = 3$) (Fig. 8C). Examining the choices of these mice on a trial-by-trial basis shows that they exhibited perseveration for ~ 10 trials and then shifted to more random behavior after ~ 15 trials (Fig. 8B). This shift can be better understood through the following analysis. Suppose that, each time a mouse made a random error, we examine the next trial where it could make the same type of error to determine whether it repeated the random error or corrected its behavior. In this way, we can compute the probability of repeating a random error in each condition; that is, control versus bilateral phasic

stimulation on incorrect trials (Fig. 8D). In control conditions, mice never repeated a random error. In stark contrast, mice that received phasic stimulation after incorrect choices during a rule change repeated random errors significantly more often—almost 50% of the time (0/10 random errors repeated in control vs 10/22 for stimulated on incorrect choice, $p = 0.013$, Fisher's exact test). In other words, phasic stimulation after incorrect trials during a rule change does not immediately elicit random errors. Rather, once mice make one random error (and receive phasic VTA–mPFC stimulation), only then do we see more random errors. We performed a similar analysis on the chance of repeating perseverative errors, but phasic stimulation after error trials did not increase the probability of repeated perseverative errors (Fig. 8E).

Bilateral tonic stimulation elicits perseveration

As part of these final experiments, we also wanted to confirm that the effect of continuous unilateral tonic stimulation (perseveration) could be elicited using bilateral stimulation. We thus repeated our original experiment using continuous tonic stimulation, but with bilateral (instead of unilateral) stimulation. When we delivered continuous bilateral tonic stimulation to mice that had learned an initial association and were learning a new rule, we found that, as before, that mice had increased difficulty learning the new rule (Fig. 8A,F and Fig. 8–1G,H, available at <https://doi.org/10.1523/JNEUROSCI.1221-17.2017.f8-1>) ($n = 8$ mice, 17.4 ± 0.7 trials in control vs 24.8 ± 2.5 trials for tonic, $p = 0.046$, Wilcoxon signed-rank test, $W = 3.5$) and made an increased number of perseverative errors (Fig. 8C) ($n = 8$ mice, 4.8 ± 0.5 perseverative errors in control, 8.9 ± 1.4 in tonic, $p = 0.02$, Wilcoxon signed-rank test, $W = 1.5$). We note that all but one mouse succeeded in reaching criterion and, examining Figure 8F, we see that mice perseverated for ~ 20 trials compared with the maximum number (30 trials) that we observed in unilaterally stimulated mice. Notably, unlike phasic stimulation, bilateral continuous tonic stimulation during a rule change did increase significantly the probability of repeating perseverative errors (20/37 perseverative errors repeated in control vs 53/70 in tonic, $p = 0.03$, Fisher's exact test) (Fig. 8E).

Discussion

Here, we report three findings about dopaminergic projections from VTA to mPFC. First, recording activity from these fibers, we observe phasic signals time locked to both rewards and cues indicating future rewards, but not reward omission. Second, tonically stimulating these fibers maintains previously learned associations. Third, phasic bursts that are not sufficient to reinforce actions in the absence of a food reward can nonetheless trigger choices that deviate from previously learned associations when they occur after choices not previously associated with reward.

Although we have not explored every possible combination of stimulation and behavior, we have endeavored to be thorough. Specifically, we tested the effects of continuous tonic or phasic VTA–mPFC stimulation on learning new associations (Fig. 5) and CPP (Fig. 3). We tested whether single phasic bursts of stimulation delivered after correct choices could reinforce learning of an initial association (Fig. 6), switching to a new association (Fig. 6), or maintenance of a previously learned association (Fig. 7). Conversely, we measured how single phasic bursts delivered after incorrect choices affect the tendency to follow a previously learned rule (Fig. 7). Finally, we evaluated how single phasic bursts delivered after correct or incorrect choices affect learning of a new association (Fig. 8). Although they do not rule out possible reinforcing effects of prefrontal dopamine under some conditions,

our results do show how prefrontal dopamine can shape behavior outside of a classic reinforcement learning framework.

Continuous stimulation models the effects of pharmacological manipulations

We began by studying continuous tonic or phasic VTA–mPFC stimulation that was not time locked to task events because we were interested in determining whether distinct patterns of stimulation could affect behavior differentially. Although we later used more naturalistic stimuli, these initial experiments are still useful because they can be thought of as the optogenetic analog of earlier experiments that have modulated dopamine receptors or levels. Consistent with these earlier studies, tonic and phasic stimulation elicited behaviors associated with the “D1” and “D2” states of the dual-state model, respectively, although we did not explore this specific pharmacology here.

Dopaminergic projections from the VTA to mPFC encode rewarding cues but are not reinforcing

Using photometry, we found that TH⁺ VTA–mPFC fibers respond similarly to other VTA dopaminergic neurons. Given that VTA neurons projecting to mPFC receive excitatory inputs from habenula encoding aversion (Lammel et al., 2012), we had expected to observe signals after aversive cues and/or reward omissions; however, we failed to find significant responses to white noise associated with a timeout before the next reward or reward omission. However, these were at best very mildly “aversive” and we have not explored more potent aversive stimuli. Moreover, single photon, bulk calcium imaging could easily miss more subtle or gradual changes in TH fiber activity; for example, activity ramps such as those observed in the striatum (Howe et al., 2013). Negative error signals may also have been more evident had we examined more dramatic shifts in reward probability such as extinction. Nonetheless, within our task parameters, mPFC-projecting dopamine neurons encode positive reward prediction errors.

It is natural to expect that the mPFC should use this reward prediction error signal to modify the animal's choices in the task to increase the chance of receiving phasic bursts in the future. Some of our observations are consistent with this possibility. However, it is difficult to reconcile the idea that mPFC dopamine signals directly reinforce specific behaviors with our findings that phasic stimulation cannot substitute for a food reward either for forming a new association or maintaining an association and that such stimulation fails to elicit CPP. One possible explanation might be to suppose that mPFC dopamine fibers are only strongly reinforcing when the mPFC is “actively engaged” in the task at hand. However, two of our observations are inconsistent with this idea. First, despite the fact that the mPFC does play an active role in switching between associations, replacing food rewards with bilateral phasic VTA–mPFC stimulation does not reinforce switching to a new association. Second, maintaining a previously learned association does not depend on the mPFC, yet phasic VTA–mPFC stimulation elicits behavior that deviates rapidly from an established association even in the absence of a rule shift. This demonstrates that dopaminergic VTA–mPFC projections can affect performance even during tasks that are classically thought of as mPFC independent.

Phasic VTA–mPFC activity can trigger behavior that deviates from previously learned rules

Although we found no evidence to support any positively (or negatively) reinforcing effects of phasic stimulation, we found

ample evidence that phasic bursts can trigger behavior that deviates from previously learned associations when stimulation was delivered either continuously or after errors (i.e., choices that were not previously associated with reward). This finding, together with the high levels of perseveration that we observed with tonic stimulation, fits well with the dual-state model discussed in the introduction. This model proposes that high levels of dopamine should be associated with enhanced flexibility and low levels with reduced flexibility. Notably, we found one important difference: phasic stimulation does not always lead to a breakdown of the current association in our experiments because pairing bilateral stimulation with a correct choice does not impair performance. Rather, we only observed deviations from the previously learned association when a phasic burst was delivered after trials on which the mouse would not expect to find a reward based on that previously learned association. This is an important distinction given that phasic increases in the activity of VTA-mPFC dopaminergic fibers seem to follow rewards; if all phasic bursts led to a breakdown of the current association, then it would be unclear how associations could ever be maintained, let alone created in the first place.

Although the effects of phasic bursts that we found are fully consistent with the dual-state model, a slightly different interpretation is that the probability of repeated an erroneous strategy is normally disrupted by dips in the dopamine signal. In this framework, delivering single phasic bursts after erroneous trials may obscure those dips, interfering with error detection and preventing the normal suppression of repeated errors, as we observed. Our experiments cannot distinguish this sort of “failure to suppress deviant responses” from direct increases in such responses.

Whereas it may be surprising that VTA-mPFC signals encode reward but are not, by themselves, positively reinforcing, we note that this result seems more natural in light of two observations. First, the PFC is not typically associated with learning of new associations, which is thought to be handled by basal ganglia circuits, where phasic dopamine release is associated with positive reinforcement (Adamantidis et al., 2011). Second, the PFC is thought to be essential for cognitive flexibility, the core aspect of which is switching to a more flexible mode of behavior when a current strategy is no longer appropriate. Our finding, that phasic bursts trigger behavior that deviates from a previously learned strategy when they follow choices that were not previously associated with reward, fits naturally within this framework. Specifically, in our task, when a mouse was repeatedly presented with two bowls, it learned to associate bowls marked by one cue with reward and bowls marked by another cue with no reward. However, the mouse will still occasionally dig in the unrewarded bowl. If the mouse unexpectedly receives a reward (or phasic burst) upon digging in a previously unrewarded bowl, it can use this new information to deviate from an outdated strategy. This represents a straightforward example of switching to a more flexible mode of behavior when a previously learned strategy is no longer appropriate. By making behavior more flexible at appropriate times, this could facilitate the ability of other brain systems (e.g., the striatum) to update cue-reward associations based on reinforcement learning mediated by dopaminergic VTA projections to basal ganglia.

Inappropriately timed phasic dopaminergic VTA-mPFC activity can elicit disorganized behavior

During a rule change, delivering phasic stimulation of VTA-mPFC projections after errors disrupted learning of the new association. Mice receiving this type of stimulation made an excess

of perseverative errors until making one random error; subsequently, they made more random errors. This experiment suggests that inappropriately timed phasic dopaminergic signals from the VTA to mPFC can elicit the sort of disorganized behavior that is a hallmark of schizophrenia and related conditions.

In summary, our findings suggest a very simple model for the effects of stimulation of the VTA-mPFC projection on behavior: tonic stimulation leads to perseveration, whereas phasic bursts can lead to behavior that deviates from previously learned rules when the bursts follow choices not previously associated with reward. These results fit naturally with dual-state model for prefrontal dopamine and with the role of the mPFC in cognitive flexibility. It will be important for future studies to determine how the effects here arise from the action of dopamine on the multiple cell types that express different dopamine receptors in the mPFC (Zheng et al., 1999; Seamans and Yang, 2004; Tseng and O'Donnell, 2007a, 2007b; Gee et al., 2012; Seong and Carter, 2012) and glutamatergic excitation in deep layers (Gorelova et al., 2012).

References

- Adamantidis AR, Tsai HC, Boutrel B, Zhang F, Stuber GD, Budygin EA, Touriño C, Bonci A, Deisseroth K, de Lecea L (2011) Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J Neurosci* 31:10829–10835. [CrossRef Medline](#)
- Adelsberger H, Garaschuk O, Konnerth A (2005) Cortical calcium waves in resting newborn mice. *Nat Neurosci* 8:988–990. [CrossRef Medline](#)
- Atasoy D, Aponte Y, Su HH, Sternson SM (2008) A FLEX switch targets Channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. *J Neurosci* 28:7025–7030. [CrossRef Medline](#)
- Bass CE, Grinevich VP, Vance ZB, Sullivan RP, Bonin KD, Budygin EA (2010) Optogenetic control of striatal dopamine release in rats. *J Neurochem* 114:1344–1352. [CrossRef Medline](#)
- Bayer HM, Glimcher PW (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47:129–141. [CrossRef Medline](#)
- Berridge KC (2007) The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl)* 191:391–431. [CrossRef Medline](#)
- Birrell JM, Brown VJ (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. *J Neurosci* 20:4320–4324. [Medline](#)
- Bissonette GB, Powell EM (2012) Reversal learning and attentional set-shifting in mice. *Neuropharmacology* 62:1168–1174. [CrossRef Medline](#)
- Bissonette GB, Martins GJ, Franz TM, Harper ES, Schoenbaum G, Powell EM (2008) Double dissociation of the effects of medial and orbital prefrontal cortical lesions on attentional and affective shifts in mice. *J Neurosci* 28:11124–11130. [CrossRef Medline](#)
- Brischoux F, Chakraborty S, Brierley DI, Ungless MA (2009) Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc Natl Acad Sci U S A* 106:4894–4899. [CrossRef Medline](#)
- Bromberg-Martin ES, Matsumoto M, Hikosaka O (2010) Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68:815–834. [CrossRef Medline](#)
- Cho KK, Hoch R, Lee AT, Patel T, Rubenstein JL, Sohal VS (2015) Gamma rhythms link prefrontal interneuron dysfunction with cognitive inflexibility in *Dlx5/6*^{+/-} mice. *Neuron* 85:1332–1343. [CrossRef Medline](#)
- Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N (2012) Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482:85–88. [CrossRef Medline](#)
- Coizet V, Dommert EJ, Redgrave P, Overton PG (2006) Nociceptive responses of midbrain dopaminergic neurons are modulated by the superior colliculus in the rat. *Neuroscience* 139:1479–1493. [CrossRef Medline](#)
- Cui G, Jun SB, Jin X, Pham MD, Vogel SS, Lovinger DM, Costa RM (2013) Concurrent activation of striatal direct and indirect pathways during action initiation. *Nature* 494:238–242. [CrossRef Medline](#)
- Cui G, Jun SB, Jin X, Luo G, Pham MD, Lovinger DM, Vogel SS, Costa RM (2014) Deep brain optical measurements of cell type-specific neural activity in behaving mice. *Nat Protoc* 9:1213–1228. [CrossRef Medline](#)
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical

- substrates for exploratory decisions in humans. *Nature* 441:876–879. [CrossRef Medline](#)
- Durstewitz D, Seamans JK (2008) The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biol Psychiatry* 64:739–749. [CrossRef Medline](#)
- Durstewitz D, Seamans JK, Sejnowski TJ (2000) Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J Neurophysiol* 83:1733–1750. [Medline](#)
- Fligel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H (2011) A selective role for dopamine in stimulus-reward learning. *Nature* 469:53–57. [CrossRef Medline](#)
- Floresco SB, Magyar O, Ghods-Sharifi S, Vexelman C, Tse MT (2006) Multiple dopamine receptor subtypes in the medial prefrontal cortex of the rat regulate set-shifting. *Neuropsychopharmacology* 31:297–309. [CrossRef Medline](#)
- Gao DM, Jeaugey L, Pollak P, Benabid AL (1990) Intensity-dependent nociceptive responses from presumed dopaminergic neurons of the substantia nigra, pars compacta in the rat and their modification by lateral habenula inputs. *Brain Res* 529:315–319. [CrossRef Medline](#)
- Gee S, Ellwood I, Patel T, Luongo F, Deisseroth K, Sohal VS (2012) Synaptic activity unmasks dopamine D2 receptor modulation of a specific class of layer V pyramidal neurons in prefrontal cortex. *J Neurosci* 32:4959–4971. [CrossRef Medline](#)
- Gorelova N, Mulholland PJ, Chandler LJ, Seamans JK (2012) The glutamatergic component of the mesocortical pathway emanating from different subregions of the ventral midbrain. *Cereb Cortex* 22:327–336. [CrossRef Medline](#)
- Grace A (1991) Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 41:1–24. [CrossRef Medline](#)
- Grace AA, Bunney BS (1983) Intracellular and extracellular electrophysiology of nigral dopaminergic neurons. 1. Identification and characterization. *Neuroscience* 10:301–315. [CrossRef Medline](#)
- Grace AA, Bunney BS (1984) The control of firing pattern in nigral dopamine neurons: burst firing. *J Neurosci* 4:2877–2890. [Medline](#)
- Gunaydin LA, Grosenick L, Finkelstein JC, Kauvar IV, Fenno LE, Adhikari A, Lammel S, Mirzabekov JJ, Airan RD, Zalocusky KA, Anikeeva P, Malenka RC, Deisseroth K (2014) Natural neural projection dynamics underlying social behavior. *Cell* 157:1535–1551. [CrossRef Medline](#)
- Howe MW, Tierney PL, Sandberg SG, Phillips PE, Graybiel AM (2013) Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* 500:575–579. [CrossRef Medline](#)
- Lammel S, Lim BK, Ran C, Huang KW, Betley MJ, Tye KM, Deisseroth K, Malenka RC (2012) Input-specific control of reward and aversion in the ventral tegmental area. *Nature* 491:212–217. [CrossRef Medline](#)
- Lapish CC, Kroener S, Durstewitz D, Lavin A, Seamans JK (2007) The ability of the mesocortical dopamine system to operate in distinct temporal modes. *Psychopharmacology (Berl)* 191:609–625. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- Mantz J, Thierry AM, Glowinski J (1989) Effect of noxious tail pinch on the discharge rate of mesocortical and mesolimbic dopamine neurons: selective activation of the mesocortical system. *Brain Res* 476:377–381. [CrossRef Medline](#)
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202. [CrossRef Medline](#)
- Overton PG, Clark D (1997) Burst firing in midbrain dopaminergic neurons. *Brain Res Rev* 25:312–334. [CrossRef Medline](#)
- Popescu AT, Zhou MR, Poo MM (2016) Phasic dopamine release in the medial prefrontal cortex enhances stimulus discrimination. *Proc Natl Acad Sci U S A* 113:E3169–E3176. [CrossRef Medline](#)
- Puig MV, Miller EK (2012) The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron* 74:874–886. [CrossRef Medline](#)
- Puig MV, Miller EK (2015) Neural substrates of dopamine D2 receptor modulated executive functions in the monkey prefrontal cortex. *Cereb Cortex* 25:2980–2987. [CrossRef Medline](#)
- Roux S, Froger C, Porsolt RD, Valverde O, Maldonado R (2003) Place preference test in rodents. *Curr Protoc Neurosci Chapter 9:Unit 9.15*. [CrossRef Medline](#)
- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115. [CrossRef Medline](#)
- Schultz W (2007) Multiple dopamine functions at different time courses. *Annu Rev Neurosci* 30:259–288. [CrossRef Medline](#)
- Schultz W (2013) Updating dopamine reward signals. *Curr Opin Neurobiol* 23:229–238. [CrossRef Medline](#)
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599. [CrossRef Medline](#)
- Seamans JK, Yang CR (2004) The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* 74:1–58. [CrossRef Medline](#)
- Seamans JK, Floresco SB, Phillips AG (1998) D1 receptor modulation of hippocampal-prefrontal cortical circuits integrating spatial memory with executive functions in the rat. *J Neurosci* 18:1613–1621. [Medline](#)
- Seong HJ, Carter AG (2012) D1 receptor modulation of action potential firing in a subpopulation of layer 5 pyramidal neurons in the prefrontal cortex. *J Neurosci* 32:10516–10521. [CrossRef Medline](#)
- Sohal VS, Huguenard JR (2003) Inhibitory interconnections control burst pattern and emergent network synchrony in reticular thalamus. *J Neurosci* 23:8978–8988. [Medline](#)
- Sohal VS, Huguenard JR (2005) Inhibitory coupling specifically generates emergent gamma oscillations in diverse cell types. *Proc Natl Acad Sci U S A* 102:18638–18643. [CrossRef Medline](#)
- Sohal VS, Zhang F, Yizhar O, Deisseroth K (2009) Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* 459:698–702. [CrossRef Medline](#)
- Stefani MR, Moghaddam B (2006) Rule learning and reward contingency are associated with dissociable patterns of dopamine activation in the rat prefrontal cortex, nucleus accumbens, and dorsal striatum. *J Neurosci* 26:8810–8818. [CrossRef Medline](#)
- 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. [CrossRef Medline](#)
- St Onge JR, Abhari H, Floresco SB (2011) Dissociable contributions by prefrontal D1 and D2 receptors to risk-based decision making. *J Neurosci* 31:8625–8633. [CrossRef Medline](#)
- Tsai HC, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K (2009) Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* 324:1080–1084. [CrossRef Medline](#)
- Tseng KY, O'Donnell P (2007b) D2 dopamine receptors recruit a GABA component for their attenuation of excitatory synaptic transmission in the adult rat prefrontal cortex. *Synapse* 61:843–850. [CrossRef Medline](#)
- Tseng KY, O'Donnell P (2007a) Dopamine modulation of prefrontal cortical interneurons changes during adolescence. *Cereb Cortex* 17:1235–1240. [Medline](#)
- Zheng P, Zhang XX, Bunney BS, Shi WX (1999) Opposite modulation of cortical N-methyl-D-aspartate receptor-mediated responses by low and high concentrations of dopamine. *Neuroscience* 91:527–535. [CrossRef Medline](#)