Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors

Chun Yun Chang1, Guillem R Esber2, Yasmin Marrero-Garcia1, Hau-Jie Yau1, Antonello Bonci1,3,4 & Geoffrey Schoenbaum1,3,5

Correlative studies have strongly linked phasic changes in dopamine activity with reward prediction error signaling. But causal evidence that these brief changes in firing actually serve as error signals to drive associative learning is more tenuous. Although there is direct evidence that brief increases can substitute for positive prediction errors, there is no comparable evidence that similarly brief pauses can substitute for negative prediction errors. In the absence of such evidence, the effect of increases in firing could reflect novelty or salience, variables also correlated with dopamine activity. Here we provide evidence in support of the proposed linkage, showing in a modified Pavlovian over-expectation task that brief pauses in the firing of dopamine neurons in rat ventral tegmental area at the time of reward are sufficient to mimic the effects of endogenous negative prediction errors. These results support the proposal that brief changes in the firing of dopamine neurons serve as full-fledged bidirectional prediction error signals.

Prediction errors—differences between predicted and actual outcomes—are thought to be responsible for associative learning. Although single-unit3–6, imaging7 and voltammetry studies8,9 have firmly established a correlative link between phasic changes in dopamine neuron activity and reward prediction error signaling, the causal evidence supporting the strong version of this proposal—that brief changes in the firing of midbrain dopamine neurons actually drive associative learning by serving as the full-fledged bidirectional prediction errors posited in learning models10—has been more tenuous (and controversial)11. Though there is direct evidence that increases in the firing of these neurons can substitute for positive prediction errors12–17, there is no comparable evidence that similarly short pauses in the activity of these notoriously slow-spiking neurons can substitute for negative prediction errors (though inhibition of midbrain indirectly via activation of projection neurons in lateral habenula at various time scales has been shown to be effective in changing behavior in a variety of settings17–20). Indeed the relatively small and very brief decrease in the firing of these neurons at the time of reward omission has led some to question whether there could be any effect on downstream targets21,22.

The bidirectional symmetry in the effects of increases and decreases in the firing of dopamine neurons is not simply the icing on the cake; it is critical to the validity of the hypothesis that these correlates are, in fact, the neural representation of these important teaching signals. Lacking such evidence, the effect of increases in firing on associative learning could be parsimoniously explained as isolated positive prediction errors or novelty or salience. Dopamine neurons have been shown to signal both novelty and salience23,24, and increases in either would be expected to facilitate—even unblock—learning25,26. Thus, demonstrating that briefly inhibiting dopamine neurons is sufficient to mimic the effects of negative prediction errors provides an acid test of the theory that dopamine neurons actually support associative learning by signaling a bidirectional prediction error signal such as that envisioned by accounts such as Rescorla-Wagner1 or temporal difference reinforcement learning2.

Here we provide such a demonstration, using as our vehicle a task called Pavlovian over-expectation. Pavlovian over-expectation is a form of extinction in which negative prediction errors are induced by heightening the expectations for reward while holding the actual reward constant. Like other forms of extinction, it shows renewal and spontaneous recovery27,28, which mark it as new learning rather than forgetting or an erasure of the old. However, unlike in conventional extinction, reward continues to be delivered and conditioned responding normally remains strong during the learning phase; indeed learning is typically only evident later in a probe test. This makes the task an excellent vehicle with which to dissociate effects of dopaminergic manipulation on learning from less specific effects that dopamine may have on vigor or motivational level29,30, attention or salience21,31 or even aversiveness32.

We modified the task to eliminate the endogenous negative prediction error by delivering the larger, expected amount of reward, and then we reintroduced these errors by briefly inhibiting tyrosine hydroxylase–positive (TH+) neurons in the ventral tegmental area (VTA) at the time of the extra reward. We found that this manipulation

1National Institute on Drug Abuse Intramural Research Program, Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, Baltimore, Maryland, USA. 2Department of Psychology, Brooklyn College, City University of New York, New York, New York, USA. 3Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University, Baltimore, Maryland, USA. 4Department of Psychiatry, The Johns Hopkins University, Baltimore, Maryland, USA. 5Department of Anatomy and Neurobiology, University of Maryland, School of Medicine, Baltimore, Maryland, USA. Correspondence should be addressed to C.Y.C. (tina.chang@nih.gov) or G.S. (geoffrey.schoenbaum@nih.gov).

Received 8 May; accepted 10 November; published online 7 December 2015; doi:10.1038/nn.4191
was sufficient to restore the extinction learning normally observed when reward is held constant. The effect was specific, inasmuch as similar inhibition delivered between trials had no effect. Inhibiting TH+ neurons also did not alter ongoing behavior, either in response to the cues or during reward consumption, suggesting that it was neither aversive nor distracting. This effect also cannot be explained by reductions in salience or associability because such changes would retard rather than promote learning.25,26 Along with prior data showing that stimulating dopamine neurons at the time of a missing positive prediction error can unblock learning,13 these results strongly support the proposal that brief phasic changes in the firing of dopamine neurons do in fact serve as bidirectional prediction errors.

RESULTS

To test the hypothesis that the suppression of firing in VTA dopamine neurons serves as a negative prediction error, we used a modified version of the Pavlovian over-expectation task33,34 consisting of three phases: conditioning, compound training and probe testing. In the conditioning phase, different cues are independently associated with food reward. Subsequently, in compound training, two of these cues are presented simultaneously to induce a heightened expectation of reward. In the standard version of the task33, this heightened expectation is violated when only the normal amount of reward is delivered, inducing a negative prediction error that modifies the strength or expression of the underlying associative representations. This change is evident as reduced responding to the individual cues when they are presented alone in the subsequent probe test. Here we modified this design by presenting the larger amount of reward predicted by the compound cue, thus eliminating the endogenous negative prediction errors. If these errors are normally signaled by brief pauses in the firing of dopamine neurons, then restoring those pauses by optogenetically inhibiting TH+ VTA neurons should restore the extinction learning observed in the normal over-expectation task.

Sixteen rats were trained in the modified Pavlovian over-expectation task (Fig. 1a). Prior to training, rats underwent surgery to infuse an opsin-encoding virus and implant optic fibers targeting the VTA (Fig. 1b and Supplementary Fig. 1). We infused AAV-Dio-NpHR3.0-eYFP (NpHR, n = 8) or AAV-Dio-eYFP (eYFP, n = 8) into the VTA of rats expressing Cre recombinase from the tyrosine hydroxylase (TH) promoter.36 Postmortem immunohistochemical verification showed a high degree of colocalization between Cre-dependent NpHR (or eYFP) and TH expression in the VTA in these rats (Fig. 1c). Quantification showed that ~85% of virus-expressing cells in the VTA (618 of 731 cells counted in sections in the anterior-posterior plane between –5.0 to –5.8 mm) were immunoreactive to anti-TH antisera. This location was chosen based on the location of the fiber implants and likely light penetration. In addition, ex vivo electrophysiology showed that both spontaneous and evoked firing in NpHR-expressing neurons in VTA were uniformly sensitive to light. In all cases, activity was immediately and reversibly silenced by 2-s pulses of green light (Fig. 1c).

Conditioning

After surgery and recovery, rats were food restricted until their body weight reached 85% of baseline, after which they started training. Training began with 12 d of conditioning, during which cues were paired with flavored sucrose pellets (banana and grape, designated as O1 and O2, counterbalanced). Three unique auditory cues (tone, white noise and clicker, designated A1, A2, and A3, counterbalanced) were the primary cues of interest. A1, the “over-expected cue”, was associated with three pellets of O1. A2, the control cue, was associated with three pellets of O2. O2 was used in order to reduce any generalization between A1 and A2. A3 was associated with no reward and thus represented the absence of a conditioned stimulus (CS-”). Rats were also trained to associate a visual cue (cue light, V) with three pellets of O1. V was to be paired with A1 in the compound phase to induce over-expectation, and therefore a nonauditory cue was used in order to discourage the compound from being perceived as a unique, distinct cue. Rats in the eYFP control and NpHR experimental groups showed similar responding to V in all phases (no main effects or interactions with group; F < 1.2, P > 0.93, Supplementary Fig. 2).

Both eYFP and NpHR rats developed elevated responding to A1 and A2, compared to A3, across the 12 sessions (Fig. 2, conditioning, dark symbols). Rats in both groups learned to respond to these two cues equally and at asymptote. In accordance with this impression, ANOVA (group × cue × session) revealed significant main effects of cue (F2,28 = 92.3, P < 0.0001) and session (F11,154 = 5.26, P < 0.0001) and a significant interaction between cue and session (F11,154 = 9.91, P < 0.0001); however, there were no main effects, nor were there any interactions with group (F < 1.8, P > 0.05). A direct comparison of

**Figure 1** Task design, fiber placements, and immunohistochemical and electrophysiological verification of Cre-dependent NpHR and eYFP expression in tyrosine hydroxylase-expressing (TH+) neurons in the VTA. (a) Top, illustration of the behavioral task. Bottom, temporal configuration of light inhibition in relative to averaged duration of pellet consumption during reward. (b) Fiber implants were localized in the vicinity of eYFP and NpHR expression in VTA. The light shading represents the maximal and the dark shading the minimal spread of expression at each level. (c) Images (left) show that majority of NpHR-expressing neurons (green) also expressed TH (red). Bottom images are expansion of the region boxed at top. Scale bar, 1 mm. Note that because the image was taken under large-field scanning, the signal intensity during acquisition was adjusted to capture the overall brightness of the entire field without ignoring relatively weak yet positive signals. This inevitably causes some area to seem be overpowered by the signal and tips the balance of color detection in the merged image, particularly in the low-magnitude image. Middle, representative traces show that NpHR-expressing neurons were responsive to light inhibition (shown as green bars). These neurons also expressed TH, as confirmed by intracellular labeling and post hoc TH staining (middle bottom). Scale bar, 50 μm. Right, spontaneous and evoked firing of NpHR-expressing neurons were interrupted by brief pulses of light inhibition (n = 10 total from three subjects; firing activity of individual neurons is summarized at the right). Error bar, s.e.m.
responding to A1 and A2 revealed no statistical effects of either cue or group or session nor any interactions during the final 4 d of conditioning ($F < 4.1, P > 0.05$).

**Compound training**

After conditioning, the rats underwent 4 d of compound training. These sessions were the same as the preceding sessions, except that V was delivered simultaneously with A1. V also continued to be presented separately, and A2 and A3 continued to be presented as before. Presentation of the A1/V compound was followed by delivery of the larger amount of reward predicted by the combined cues, that is, six pellets—the three predicted by A1 plus the three predicted by V. However, the rats also received three pulses of light via the optical fibers implanted in VTA. In the “reward” run, this light pattern was delivered during the second half of food pellet consumption (Fig. 1a), timed to mimic presumed changes in dopamine neuron firing at the time of omission of reward in our standard over-expectation task (see Online Methods). The duration of each light pulse approximated the duration of inhibited firing at the time of omission of reward in our dopamine neuron firing at the time of omission of reward in our compound training (middle column) during the subsequent probe tests (right column).

Conditioned responding is represented as the percentage of time the rats spent in the food cup during the cues. Percentage of time rats spent in the food cup during the food consumption period during compound training, when TH+ neurons were inhibited, is shown in the bar graph insets. Data from the reward run (first and third rows, Reward), when TH+ neurons were inhibited at the time of reward, are shown with dark symbols; data from the ITI run (second and fourth rows, ITI), when TH+ neurons were inhibited in the ITI, are shown with light symbols. $n = 8$ for each group. Vertical bars show s.e.m. NS, nonsignificant at $P > 0.1$; *$P < 0.0001$ based on stats given in main text. All values in each line plot represent the percentage of time rats spent in the food cup during the CS after correction for rearing.

Thus light-induced inhibition of TH+ neurons in VTA, either during food delivery or later during the intertrial interval, had no effect on established Pavlovian conditioned responding. There was also no impact of light-induced inhibition of TH+ neurons on time spent in the food cup after food delivery (Fig. 2, compound training, inset bar graphs); rats spent more time in the food cup after the compound cue, reflecting the larger amount of reward delivered in these trials (main effect of cue: $F_{1,28} = 196.3, P < 0.0001$). However, there were no main effects or interactions in these measures involving either group or run ($F < 3.1, P > 0.09$), and all rats ate all the food pellets available in every session (food cups were inspected at the conclusion of each session and also during three randomly selected intertrial intervals for each rat during compound training). The similarity in all of these behavioral measures across groups and also across runs within the NpHR group is strong evidence that brief inhibition of TH+ VTA neurons at the time of reward was neither distracting nor aversive.

**Probe testing**

But did inhibition of the TH+ neurons in VTA affect the strength of the underlying associative representations?

To address this question, the rats received a probe test after the completion of each run of compound conditioning, in which A1, A2 and A3 were presented alone without reinforcement. In probe tests at the end of both the reward (Fig. 2, probe tests, dark symbols) and ITI runs
**DISCUSSION**

In this study we used a modified version of a Pavlovian over-expectation task, in which the heightened reward expectations were met, to probe the sufficiency of presumed negative error signals from dopamine neurons to support extinction learning. We found that brief optogenetic inhibition of TH+ neurons in the VTA, designed to mimic the negative prediction errors signaled by midbrain dopaminergic neurons, was sufficient to restore the extinction learning normally driven by over-expectation in this task.\(^{13,34}\) The optogenetically driven extinction learning was specific inasmuch as it was observed in NpHR rats but not in eYFP controls. Additionally, learning was observed only when TH+ neurons were inhibited around the time of expected reward; there was no learning when the same neurons were inhibited during the intertrial intervals. Although we did not counterbalance the order in which the reward versus intertrial interval inhibition was given, we have previously found that we can retrain in this manner, re-establish normal responding, and then reproduce the over-expectation effect a second and in some cases a third time.\(^{44}\) Indeed we have found that the effect is somewhat stronger in reiterations of the procedure, yet in the current experiment, despite the fact that the ITI inhibition happened second, the NpHR group showed no trace of an effect.

The inhibition of the TH+ neurons also had no direct effect in the sessions in which it occurred. And we never found unate pellets. The effect of inhibition was apparent only during the subsequent probe test.
Thus, inhibiting TH+ neurons at the time of reward was not obviously aversive or distracting, nor did it directly affect the palatability of the reward or the overall vigor of responding or motivational level in the task, as all of these effects would be expected to disrupt behaviors immediately. It is also not well explained by reduced salience or attention to the learning materials, because a decline in salience would affect performance and retard rather than promote learning. Instead, this pattern of results suggests that brief pauses in dopamine neuron firing drive extinction learning because they act as negative prediction errors, causing relatively subtle but specific learning. Along with prior data showing that stimulating dopamine neurons at the time of a missing positive prediction error can restore blocked learning, these results support the proposal that brief phasic changes in the firing of dopamine neurons, similar to what happens when unexpected rewards are encountered or expected rewards are omitted, serve as bidirectional prediction error signals to drive associative learning.

Our study (and the most relevant prior study) used TH-Cre rats. Recently, the use of transgenic Cre-driver lines to selectively target midbrain dopamine neurons has come under fire due to reports that viral expression in TH-Cre mice is often poorly restricted to TH-immunoreactive neurons. Although the results were regionally heterogeneous, the co-localization of viral and TH protein was as low as 50% in some parts of the midbrain. This may occur because it takes only a small amount of Cre to allow viral insertion and subsequent expression, whereas it requires substantial amounts of TH protein for detection by conventional immunostaining. Morales and colleagues have reported that many neurons in mouse VTA may suffer from this problem. However, they did not find this dissociation in rat VTA, where neurons expressing TH mRNA were uniformly immunoreactive for TH protein as well as for aromatic L-amino acid decarboxylase (AADC), another enzyme necessary for dopamine synthesis, suggesting that studies using TH-Cre rats are less likely to erroneously ascribe the effects of optogenetic manipulations to neurons incapable of synthesizing dopamine.

A separate issue is that some TH+ neurons in VTA do not express vesicular monoamine transporter 2 (VMAT2), an enzyme necessary for packaging for the vesicular release of dopamine. This appears to be true across species. However, these TH+VMAT2− neurons tend to be located more medially in areas projecting to midline regions such as habenula. We believe this concern is therefore somewhat mitigated in our experiment because our infusions and fiber placements were targeted to more laterally within VTA, to locations where TH+ neurons are more likely to have VMAT2 (ref. 41).

Of course, it is now well established that dopamine neurons often co-release other factors, such as glutamate and GABA, suggesting that studies using TH-Cre rats are less likely to be due to suppression of the release of these other neurotransmitters downstream, either alone or in combination with the suppression of dopamine release.

Our data are consistent with the effects of inhibiting midbrain indirectly via activation of projection neurons in lateral habenula. Lateral habenula neurons have been shown to fire in response to negative prediction errors and to exert an inhibitory influence on dopamine neurons in the VTA. Thus, activating them in vivo is thought to cause suppression of dopamine neuron firing. Most relevant to the current study, stimulation of lateral habenula electrically or optogenetically has been found to reduce responding directed toward, or choices of, reward-predictive cues when given at the time of cue presentation and to reduce subsequent choices of a particular option when given at the time of reward. In these cases, stimulation was temporally specific and had effects on reward-related behavior that are consistent with signaling of negative prediction errors.

We now show that similar effects can be produced by direct inhibition of TH+ neurons in the rat VTA. We can also rule out a number of alternative explanations for the effect we observed, such as changes in salience or in attention or aversive effects of inhibiting the dopamine neurons. These variables have been shown to correlate with and to be causally related to dopamine neuron firing, but they are not viable explanations for our observations.

Beyond supporting the prediction error hypothesis, these data offer further insight into the role of dopamine. For example, it is notable that suppressing the relatively low baseline activity of dopamine neurons was sufficient to drive learning. In the past, the plausibility of downstream effects from such a small change in firing has been questioned, leading to suggestions that long-duration pauses or other manipulations might be necessary for learning in response to negative prediction errors. Although our results do not exclude these possibilities, they do show that brief changes in the firing of VTA dopamine neurons are sufficient to serve as error signals. It would be of interest to explore in detail whether longer pauses or other parametric manipulations would be more effective at inducing learning in this setting, especially as the behavioral effect found here is modest. There is evidence, however, that the relatively small decreases in spiking activity induced by negative prediction errors and mimicked here by our experimental manipulation may translate into much larger declines in dopamine efflux in terminal regions, perhaps even matching the increases in efflux seen in response to positive prediction errors. These data provide a mechanism whereby brief pauses in baseline firing may be sufficient to produce outsized downstream effects. It is worth noting that the effect here is similar to what we normally observe in Pavlovian over-expectation when the larger amount of reward is not delivered, suggesting that we are reproducing the full effect of the negative prediction errors we are trying to mimic.

It is also worth noting that the effects reported here differ from those we reported previously after pharmacological inhibition of VTA via infusion of a GABA agonist cocktail. In that study, we found that inactivation of VTA during compound training disrupted learning in the standard over-expectation task. We had speculated that the effect on summation might reflect tonic changes in dopamine, but that the learning deficit was likely to be due to a loss of phasic error signals, since the neurons were likely unable to suppress firing further upon reward omission. The results of the current study support this contention inasmuch as we were able to reinstate extinction learning with brief pauses in dopaminergic activity without affecting cue-evoked or post-cue responding.

Finally, a third feature worth noting is that our effect was cue specific. That is, although our manipulation was general (that is, not designed to reproduce any special pattern or ensemble response), it affected only the associative strength of the cues that predicted the event with which it was paired. Responding to the other, very similar auditory cues was not affected, and when we applied it even a minute later, during the intertrial interval, even this effect disappeared. Further, by reinforcing the visual cue on separate trials, we were able to counteract the effects of its pairing with the inhibition. This was true even though the visual cue predicted the same reward, delivered with the same timing and in the same location, as the auditory cue that was extinguished. Thus the ‘credit’ for the induced pauses in firing of the TH+ neurons was assigned specifically and appropriately to the cue that would have received that credit had the additional reward actually not been delivered. This indicates that short pauses in the firing of TH+ VTA neurons, as a population, act precisely as predicted for a negative prediction error both generally, in terms of extinguishing responding to Pavlovian cues that are nearby in time, and more specifically in terms of where responsibility for the negative prediction error is assigned.
The current results suggest that brief pauses in dopamine neuron activity that mirror phasic changes in dopamine firing are sufficient to reinstate a relatively complex and specific form of extinction learning. This observation, along with analogous evidence that phasic increases in dopamine activity can reinstate learning in the face of blocking, provides strong causal evidence that these phasic changes function as reward prediction errors.

METHODS

Methods and any associated references are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program at the US National Institute on Drug Abuse (NIDA). The authors would like to thank K. Deisseroth and the Gene Therapy Center at the University of North Carolina at Chapel Hill for providing viral reagents, and G. Stuber for technical advice on their use. We would also like to thank B. Harvey and the NIDA Optogenetic and Transgenic Core and M. Morales and the NIDA Histology Core for their assistance. The authors would like to thank K. Deisseroth for providing viral reagents, and G. Stuber for technical advice on their use. We would also like to thank B. Harvey and the NIDA Optogenetic and Transgenic Core for their assistance. The opinions expressed in this article are the authors’ own and do not reflect the views of the US National Institutes of Health/Department of Health and Human Services.

AUTHOR CONTRIBUTIONS

C.Y.C. and G.S. conceived the experiment; C.Y.C. carried out the experiment, with help from G.R.E. and Y.M.-G. on the behavioral design and histology and from H.-J.Y. and A.B. on the slice physiology; C.Y.C. and G.S. analyzed the data and prepared the manuscript, in consultation with the other authors, particularly G.R.E., whose input on learning theory issues was invaluable.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Rescator, R.A. & Wagner, A.R. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. in Classical Conditioning II: Current Research and Theory (eds. Black, A.H., & Prokasy, W.F.) 69–99 (Appleton-Century-Crofts, New York, 1972).

2. Sutton, R.S. Learning to predict by the method of temporal difference. Mach. Learn. 9, 9–44 (1988).

3. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. J. Neurophysiol. 72, 1024–1027 (1994).

4. Roesch, M.R., Calu, D.J. & Schounbaum, G. Dopamine neurons encode the better option in rats during between-differently delayed or sized rewards. Nat. Neurosci. 10, 1615–1624 (2007).

5. Waelchli, P., Dickinson, M. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. Nature 412, 43–48 (2001).

6. Pan, W.-X., Schmidt, R., Wickers, J.R. & Hyland, B.I. Dopamine cells respond to help from G.R.E. and Y.M.-G. on the behavioral design and histology and from H.-J.Y. and A.B. on the slice physiology; C.Y.C. and G.S. analyzed the data and prepared the manuscript, in consultation with the other authors, particularly G.R.E., whose input on learning theory issues was invaluable.

The current results suggest that brief pauses in dopamine neuron activity that mirror phasic changes in dopamine firing are sufficient to reinstate a relatively complex and specific form of extinction learning. This observation, along with analogous evidence that phasic increases in dopamine activity can reinstate learning in the face of blocking, provides strong causal evidence that these phasic changes function as reward prediction errors.
**ONLINE METHODS**

**Subjects.** Eighteen male transgenic rats that carried a TH-dependent Cre expression system in a Long-Evans background (NIDA animal breeding facility) were used in this study. The rats were housed in a 12 h light/12 h dark cycle with unlimited access to food and water, except during the behavioral experiment, when they were food restricted to maintain 85% of their baseline weight. All experimental procedures were conducted in accordance with Institutional Animal Care and Use Committee of the US National Institute of Health guidelines.

**Surgical procedures.** Rats (>275 g) received bilateral infusions of AAV5-EF1α-DIO-NP-HR3.0-eYFP (5 males and 3 females) or AAV5-EF1α-DIO-eYFP (6 males and 2 females) into the VTA (anterior-posterior (AP): −5.3 mm (referred to bregma), mediolateral (ML): ±0.7 mm (referred to the midline), and dorsoventral (DV): −7.0 mm and −8.2 mm for male and −6.7 mm and −7.9 mm for female (referred to the brain surface). Virus was obtained from the University of North Carolina at Chapel Hill Gene Therapy Center, courtesy of K. Deisseroth. A total of 1–1.5 µl of virus with a titer of ≥10^{12} virus genomes (vg) was injected at the rate of 0.1 µl/min per injection site. The rats also were implanted with optic fibers bilaterally (200 µm diameter, Thorlab, NJ; AP: −5.3 mm; ML: ±2.61 mm; and DV: −7.5 mm for male and −7.2 mm for female at 15° angle pointing to the midline).

**Apparatus.** Training was conducted in 8 standard behavioral chambers from Coulbourn Instruments (Allentown, PA), each enclosed in a sound-resistant shell. A food cup was recessed in the center of one end wall. Entries were monitored by a photobeam. A food dispenser containing 45 mg sucrose pellets (plain, banana-flavored, or grape-flavored; Bio-serv, Flemington, NJ) allowed delivery of pellets into the food cup. The pellets were delivered at a rate of 1 pellet per 0.8 s, beginning immediately after cue presentation. White noise or a tone, each measuring approximately 76 dB, was delivered via a wall speaker. Also mounted on that wall was a clicker (−1 Hz) and a 6-W bulb that could be illuminated to provide a light stimulus during the otherwise dark session.

**Optogenetically driven Pavlovian over-expectation.** Rats were shaped to retrieve food pellets, then they underwent 12 conditioning sessions. In each session, the rats received eight 30-s presentations of three different auditory stimuli (A1, A2, and A3) and one visual stimulus (V), in a blocked design in which the order of cue-blocks was counterbalanced. Intertrial intervals (ITIs) were 150 s. For all conditioning, V consisted of a cue light, and A1, A2, and A3 consisted of a tone, a clicker, or white noise (counterbalanced). Two differently flavored sucrose pellets (banana and grape, designated O1 and O2, counterbalanced) were used as rewards. V and A1 terminated with delivery of three pellets of O2, and A2 terminated with delivery of six pellets of O1. A3 was followed by a 1.5-s interval. Of course this reflects an assumption on our part; in reality the timing of negative prediction errors in this setting is likely quite variable, dependent on the rat behavior. However, the timing of our light delivery has the advantage of not interfering with the perception of the normally delivered pellets, thereby reducing the chances of effects due to changes in the reinforcing effect of the normally delivered pellets.

**Histology and immunohistochemistry.** Rats that received viral infusions and fiber implants were euthanized with an overdose of isoflurane and perfused with 1× phosphate-buffered saline (PBS) followed by 4% paraformaldehyde (Santa Cruz Biotechnology Inc., CA). Fixed brains were cut in 40 µm sections to examine fiber tip position under fluorescence microscope (Olympus Microscopy, Japan). For immunohistochemistry, the brain slices were first blocked in 10% goat serum made in 0.1% Triton-X100/1× PBS and then incubated in anti-tyrosine hydroxylase (TH) antisera (MAB318, 1:600; EMD Millipore, Billerica, Massachusetts) followed by Alexa 568 secondary antisera (A11031, 1:1,000, Invitrogen, Carlsbad, CA). The image of brain slices were acquired by fluorescence Virtual Slide microscope (Olympus America, Melville, NY), and later analyzed in Adobe Photoshop. The VTA, including anterior (rostral and parabrachial pigmentation area) and posterior (caudal, parabrachial pigmentation area, paranigral nucleus, and medial substantia nigra pars medialis), of brain slices from AP−5.0 mm to −5.8 mm from 5 subjects were analyzed. This encompasses the location targeted by our fibers and likely to achieve good light penetration. For quantification, the intensity of 4 random 40 µm × 40 µm square areas from the background were averaged to provide a baseline, and positive staining was defined as signal 2.5 times this baseline intensity, with a cell diameter larger than 5 µm, co-localized within cells reactive to DAPI staining.

**Ex vivo electrophysiology.** Three additional TH-Cre rats that received AAV-EF1α-DIO-NP-HR3.0-eYFP injections in the VTA were anesthetized with isoflurane and perfused with 4 ml ice-cold NMDG-based artificial CSF (aCSF) solution containing (in mM): 92 NMDG, 20 HEPES, 2.5 KCl, 1.2 NaH2PO4, 10 MgSO4, 0.5 CaCl2, 30 NaHCO3, 25 glucose, 2 thiourea, 5 Na-aspartate, 3 Na-pyruvate, and 12 N-acetyl-L-cysteine (300–310 µm, pH 7.3–7.4). After perfusion, the brains were immediately removed and the horizontal brain slices containing the VTA were made in 300 µm thick slices (Leica, Nussloch, Germany). The brain slices were recovered within 12 min at 32 °C in NMDG-based aCSF and incubated for at least 1 h in HEPES-based aCSF containing (in mM): 92 NaCl, 20 HEPES, 2.5 KCl, 1.2 NaH2PO4, 1 MgSO4, 2 CaCl2, 30 NaHCO3, 25 glucose, 2 thiourea, 5 Na-aspartate, 3 Na-pyruvate, and 12 N-acetyl-L-cysteine (300–310 µm, pH 7.3–7.4, room temperature). During the recording, the brain slices were superfused with standard aCSF constituted (in mM) of 125 NaCl, 2.5 KCl, 1.25 NaH2PO4, 1 MgCl2, 2.4 CaCl2, 26 NaHCO3, 11 glucose, 0.1 picrotoxin, and 2 kynurenic acid, and saturated with 95% O2 and 5% CO2 at 32 °C–34 °C. Glass pipette (pipette resistance around 2–2.5 MΩ, King Precision Glass, Claremont, CA) filled with K+ based internal solution (in mM: 140 KMeSO4, 5 KCl, 0.05 EGTA, 2 MgCl2, 2 Na3ATP, 0.4 NaGTP, 10 HEPES, and 0.05 Alexa Fluor 594 (Invitrogen), pH 7.3, 290 mOsm) was used throughout the experiment. Cell-attached and whole-cell configurations were made from identified eYFP* cells using MultiClamp 700B amplifier (Molecular Devices, Sunnyvale, CA). To verify the functional expression of nPHR in the recorded cells, a 2-s long pulse of green light (532 nm) were delivered at
the intensity of 0.4–0.8 mW via an optic fiber that was attached to a 40x objective lens and positioned right above the slice. NpHR expression was confirmed by a membrane hyperpolarization under current clamp, or an outward current under voltage clamp upon light stimulation. To confirm the efficiency of short light pulses in inhibiting neuronal activity, pulses of green light (2 s in duration with 1.5 s interval) were delivered while neuronal firing was evoked by a 10-s train of somatic current injection (0.5–1.5 nA for 2 ms at 5 Hz). After the recording, some of the cells were filled with dye intracellularly for post hoc immunohistochemical verification of TH expression. Throughout the recording, series resistance (10–15 MΩ) was continually monitored online with a 20 pA, 300 ms current injection after every current injection step. The cell was excluded if the series resistance changed for more than 20%. Signal was sampled at 20 kHz and filtered at 10 kHz. Data were acquired in Clampex 10.3 (Molecular Devices, Foster City, CA), and were analyzed off-line in Clampfit 10.5 (Molecular Devices) and IGOR Pro 6.3 (WaveMetrics, Lake Oswego, OR). For post hoc staining, brain slices were fixed in 4% PFA, washed with 2% Triton-X100 (1 h), blocked in 3% normal donkey serum and stained with mouse anti-TH antisera (T2928; 1:1,500, Sigma-Aldrich, St. Louis, MO). After immunoreaction to DyLight 405-conjugated secondary antibody (715-475-150; 1:500, Jackson ImmunoResearch, West Grove, PA), the slices were mounted with Mowiol mounting solution (Sigma-Aldrich) and examined under confocal microscope (Fluoview FV1000, Olympus).

A Supplementary Methods Checklist is available.