

1-2016

# Brief Optogenetic Inhibition of Dopamine Neurons Mimics Endogenous Negative Reward Prediction Errors

Chun Yun Chang

*National Institute on Drug Abuse Intramural Research Program*

Guillem R. Esber

*CUNY Brooklyn College*

Yasmin Marrero-Garcia

*National Institute on Drug Abuse Intramural Research Program*

Hau-Jie Yau

*National Institute on Drug Abuse Intramural Research Program*

Antonello Bonci

*Johns Hopkins University*

*See next page for additional authors*

**[How does access to this work benefit you? Let us know!](#)**

Follow this and additional works at: [https://academicworks.cuny.edu/bc\\_pubs](https://academicworks.cuny.edu/bc_pubs)

 Part of the [Psychiatry and Psychology Commons](#)

---

## Recommended Citation

Chang, Chun Yun; Esber, Guillem R.; Marrero-Garcia, Yasmin; Yau, Hau-Jie; Bonci, Antonello; and Schoenbaum, Geoffrey, "Brief Optogenetic Inhibition of Dopamine Neurons Mimics Endogenous Negative Reward Prediction Errors" (2016). *CUNY Academic Works*.

[https://academicworks.cuny.edu/bc\\_pubs/139](https://academicworks.cuny.edu/bc_pubs/139)

This Article is brought to you for free and open access by the Brooklyn College at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu).

---

**Authors**

Chun Yun Chang, Guillem R. Esber, Yasmin Marrero-Garcia, Hau-Jie Yau, Antonello Bonci, and Geoffrey Schoenbaum



# HHS Public Access

Author manuscript

*Nat Neurosci.* Author manuscript; available in PMC 2016 June 07.

Published in final edited form as:

*Nat Neurosci.* 2016 January ; 19(1): 111–116. doi:10.1038/nn.4191.

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

Chun Yun Chang<sup>1,\*</sup>, Guillem R Esber<sup>5</sup>, Yasmin Marrero-Garcia<sup>1</sup>, Hau-Jie Yau<sup>1</sup>, Antonello Bonci<sup>1,3,4</sup>, and Geoffrey Schoenbaum<sup>1,2,3,\*</sup>

<sup>1</sup>National Institute on Drug Abuse Intramural Research Program, Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, Baltimore, MD 21224

<sup>2</sup>Department of Anatomy and Neurobiology, University of Maryland, School of Medicine, Baltimore, MD 21201

<sup>3</sup>Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University, Baltimore, MD 21287

<sup>4</sup>Department of Psychiatry, The Johns Hopkins University, Baltimore, MD 21287

<sup>5</sup>Brooklyn College, City University of New York, New York, NY 11210

### Abstract

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. While 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. Lacking such evidence, the effect of increases in firing could reflect novelty or salience, variables also correlated with dopamine activity. Here we provide such evidence, 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<sup>1, 2</sup>. While single unit<sup>3–6</sup>, imaging<sup>7</sup>, and voltametry studies<sup>8, 9</sup> have firmly established a correlative link between phasic changes in dopamine neuron activity and reward prediction error signaling, the causal evidence supporting the

---

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

Correspondence or requests for material should be addressed to C.Y.C. (; Email: [tina.chang@nih.gov](mailto:tina.chang@nih.gov)) or G.S. (; Email: [geoffrey.schoenbaum@nih.gov](mailto:geoffrey.schoenbaum@nih.gov))

#### 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.

The authors declare that they have no competing financial interests or conflicts related to the data presented in this manuscript.

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 models<sup>10</sup> – is more tenuous (and controversial)<sup>11</sup>. While there is direct evidence that increases in the firing of these neurons can substitute for positive prediction errors<sup>12-17</sup>, 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 timescales has been shown to be effective in changing behavior in a variety of settings<sup>17-20</sup>). 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 targets<sup>21, 22</sup>.

The bidirectional symmetry in the effect 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 as novelty or salience. Dopamine neurons have been shown to signal both novelty and salience<sup>23, 24</sup>, and increases in either would be expected to facilitate – even unblock – learning<sup>25, 26</sup>. 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 like that envisioned by accounts such as Rescorla-Wagner<sup>1</sup> or temporal difference reinforcement learning<sup>2</sup>.

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 recovery<sup>27, 28</sup>, which mark it as new learning rather than forgetting or an erasure of the old. However, unlike 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 dopamine may have on vigor or motivational level<sup>29, 30</sup>, attention or salience<sup>11, 31</sup>, or even aversiveness<sup>32</sup>.

We modified the task to eliminate the endogenous negative prediction error by delivering the larger, expected amount of reward, then we re-introduced these errors by briefly inhibiting tyrosine hydroxylase positive (TH+) neurons in VTA at the time of the extra reward. We found that this manipulation 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 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, since such changes would retard rather than promote learning<sup>25, 26</sup>. Along with prior data showing that stimulating dopamine neurons at the time of a missing

positive prediction error can unblock learning<sup>13</sup>, 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 suppression of firing in VTA dopamine neurons serves as a negative prediction error, we used a modified version of the Pavlovian over-expectation task<sup>33, 34</sup> 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 task<sup>35</sup>, 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 Fig. S1). 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<sup>36</sup>. Post-mortem immunohistochemical verification showed a high degree of co-localization between Cre-dependent NpHR (or eYFP) and TH expression in 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 was uniformly sensitive to light. In all cases, activity was immediately and reversibly silenced by 2-second pulses of green light (Fig. 1C).

## Conditioning

After surgery and recovery, rats were food restricted until their body weight reached to 85% of baseline, after which they started training. Training began with 12 days 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 served as a 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; therefore a non-auditory 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;  $F$ 's  $< 1.2$ ,  $p$ 's  $> 0.93$ ., Fig S2).

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 accord with this impression, ANOVA (group  $\times$  cue  $\times$  session) revealed significant main effects of cue ( $F_{(2,28)}=92.3$ ,  $p<0.0001$ ) and session ( $F_{(11,154)}=5.26$ ,  $p<0.0001$ ), and a significant interaction between cue and session ( $F_{(22,308)}=9.91$ ,  $p<0.0001$ ); however there were no main effects nor any interactions with group ( $F$ 's  $< 1.8$ ,  $p$ 's  $> 0.05$ ). A direct comparison of responding to A1 and A2 revealed no statistical effects of either cue or group or session nor any interactions during the final 4 days of conditioning ( $F$ 's  $< 4.1$ ,  $p$ 's  $> 0.05$ ).

### Compound training

After conditioning, the rats underwent four days of compound training. These sessions were the same as 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, i.e. 6 pellets - the three predicted by A1 plus the 3 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 Methods). The duration of each light pulse approximated the duration of inhibited firing typically observed in dopamine neurons upon omission of an expected reward<sup>3-6</sup>. In the subsequent "ITI" control run, conducted after reminder training to restore responding on the original associations, the same light pattern was delivered 75-s into the 150-s inter-trial interval (Fig 1A), as a control for non-specific effects.

Both eYFP and NpHR rats maintained elevated responding to A1/V and A2, compared to A3, across the four sessions in both the reward (Fig. 2, Compound Training, dark symbols) and ITI runs (Fig. 2, Compound Training, light symbols). ANOVA's (group  $\times$  cue  $\times$  session) revealed significant main effects of cue (reward run:  $F_{(2,28)}=155.3$ ,  $p<0.0001$ ; ITI run:  $F_{(2,28)}=70.1$ ,  $p<0.0001$ ) and session (reward run:  $F_{(3,42)}=3.28$ ,  $p=0.03$ ), but no main effects of nor any interactions with group in either run ( $F$ 's  $< 1.9$ ,  $p$ 's  $> 0.1$ ), and separate analyses of responding to each cue in each run found no group effects ( $F$ 's  $< 2.3$ ,  $p$ 's  $> 0.1$ ). Thus light-induced inhibition of TH+ neurons in VTA, either during food delivery or later during the inter-trial interval, had no effect on established Pavlovian conditioned responding. Nor was there any 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 on 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 ( $F$ 's < 3.1,  $p$ 's > 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 ITI's 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 (Fig. 2, Probe Tests, light symbols), the rats showed elevated responding to A1 and A2, compared to A3, and this responding extinguished across each session. ANOVA's (group  $\times$  cue  $\times$  trial) revealed significant main effects of cue and trial (reward: cue:  $F_{(2,28)}=82.5$ ,  $p<0.0001$ ; trial:  $F_{(7,98)}=20.8$ ,  $p<0.0001$ ; ITI: cue:  $F_{(2,28)}=25.9$ ,  $p<0.0001$ ; trial:  $F_{(7,98)}=23.7$ ,  $p<0.0001$ ) and significant interactions between cue and trial (reward:  $F_{(14,196)}=6.19$ ,  $p<0.0001$ ; ITI:  $F_{(14,196)}=6.39$ ,  $p<0.0001$ ). Importantly there were no significant main effects of group nor any group by trial interactions in either run ( $F$ 's < 2.2,  $p$ 's > 0.15), indicating that there were no effects of prior light delivery on general responding or extinction learning within the probe test.

However, in addition to these effects, which were similar in both runs, the NpHR rats also showed less responding to the A1 "over-expected" cue than to the A2 control cue in the probe test conducted after inhibition of the TH+ neurons during reward (Fig. 2, Probe Tests, dark symbols). This difference was not generally present in the responding of the control rats, although their behavior to A1 did decline on the last 2 trials of the session. As a result, there was a significant interaction between group, cue, and trial ( $F_{(7,98)}=2.16$ ,  $p=0.04$ ), and a step-down ANOVA comparing responding to A1 and A2 revealed that NpHR rats showed less responding to the over-expected A1 cue than to the A2 control cue throughout the probe test ( $F_{(1,14)}=22.2$ ,  $p=0.0003$ ; Fig. 2b, Probe Tests, dark symbols). By contrast, the eYFP rats responded at identical levels to the A1 and A2 cues until the final 2 trials ( $F_{(1, 14)}=2.8$ ,  $p=0.12$ ; Fig. 2a, Probe Tests, dark symbols). Notably this difference in responding was not observed in the probe test conducted after inhibition of the TH+ neurons during the ITI (Fig 2, Probe Tests, light symbols). Both NpHR and eYFP rats showed similar responding to A1 and A2, and there was neither a main effect nor any interactions involving group ( $F$ 's < 1.6,  $p$ 's > 0.10). A direct comparison of responding to A1 and A2 in the two probe tests found a significant interaction between cue and probe test in the NpHR rats ( $F_{(1, 14)}=7.72$ ,  $p=0.01$ ), consistent with a response difference after inhibition during reward ( $p<0.0001$ ) but not after inhibition during the ITI ( $p=0.85$ ). By contrast, there was no interaction in the data from the eYFP controls ( $F_{(1, 14)}=2.45$ ,  $p=0.14$ ; all other effects involving probe test also N.S.:  $F$ 's < 1.1,  $p$ 's > 0.3), nor was there any main effect or any interaction with cue when both probe tests were considered ( $F$ 's < 2.5,  $p$ 's > 0.1).

The impact of inhibition during reward was also evident in a comparison of the difference in A2 and A1 responding within individual rats in the two probe tests (Fig. 3). Despite having relatively few subjects in each group, the difference in these scores was significantly larger in the probe test than it was before compound training only in the NpHR group and only when light was delivered during reward (Fig. 3b, dark histogram). Importantly this did not reflect any difference in responding to A2 ( $F$ 's < 2.59,  $p$ 's > 0.16). There was no change in these difference scores when the same light pattern was delivered in the ITI (Fig. 3b, light histogram) or when it was delivered to rats in the eYFP group during reward (Fig. 3a, dark histogram) or during the ITI (Fig. 3a, light histogram).

## 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 supporting 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<sup>33, 34</sup>. The optogenetically-driven extinction learning was specific inasmuch as it was observed in NpHR rats but not in eYFP controls. Additionally, learning was only observed when TH+ neurons were inhibited around the time of expected reward delivery; 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 inter-trial interval inhibition was given, we have previously found that we can retrain in this manner, reestablish normal responding, and then reproduce the over-expectation effect a second and in some cases a third time<sup>34</sup>. 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 uneaten 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 overall vigor of responding or motivational level in the task, since 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, since a decline in salience would affect performance and retard rather than promote learning<sup>25, 26</sup>. 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<sup>33, 34</sup>. Along with prior data showing that stimulating dopamine neurons at the time of a missing positive prediction error can restore blocked learning<sup>13</sup>, 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<sup>10</sup>.

Our study (and the most relevant prior study<sup>13</sup>) used TH-Cre rats<sup>36</sup>. 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<sup>37,38</sup>. While the results were regionally heterogeneous, the colocalization 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<sup>39</sup>. Morales and colleagues have reported that many neurons in mouse VTA may suffer from this problem<sup>40</sup>. 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<sup>41</sup>, 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<sup>42</sup>. However these TH+/VMAT2- neurons tend to be located more medially<sup>42</sup>, in areas projecting to midline regions such as habenula. We believe this concern is therefore somewhat mitigated in our experiment, since our infusions and fiber placements were targeted more laterally within VTA, where TH+ neurons are more likely to have VMAT2<sup>41</sup>.

Of course, it is now well established that dopamine neurons often co-release other factors, such as glutamate and GABA<sup>41,43-45</sup>. Our effects might 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<sup>17,20</sup>. Lateral habenula neurons have been shown to fire in response to negative prediction errors<sup>46</sup> and to exert an inhibitory influence on dopamine neurons in the VTA<sup>46-48</sup>. Thus activating them in vivo is thought to cause suppression of dopamine neuron firing. Most relevant to the current study, stimulating lateral habenula electrically or optogenetically has been found to reduce responding directed towards or choices of reward predictive cues when given at the time of cue presentation<sup>17,20</sup> and to reduce subsequent choices of a particular option when given at the time of reward<sup>17</sup>. 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, attention, or aversive effects of inhibiting the dopamine neurons. These variables have been shown to correlate with<sup>23,24</sup> and to be causally related to<sup>18,19,49,50</sup> 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 cause 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 neurotransmitter systems might be necessary for learning in response to negative prediction errors<sup>21, 22</sup>. While our results do not exclude these possibilities, they do show that brief changes in the firing of VTA dopamine neurons are sufficient. 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. However there is evidence that the relatively small declines 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<sup>9</sup>. 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<sup>33, 34</sup>, 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<sup>33</sup>. 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 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 (i.e. not designed to reproduce any special pattern or ensemble response), it only affected 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 been not 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.

## Online Methods

### Subjects

Eighteen male transgenic rats that carried TH-dependent Cre expressing system on a Long-Evans background (NIDA animal breeding facility) were used in this study<sup>36</sup>. The rats were maintained in a 12-h light/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-EF1 $\alpha$ -DIO-NpHR3.0-eYFP (5 males and 3 females) or AAV5-EF1 $\alpha$ -DIO-eYFP (6 males and 2 females) into the VTA (AP: -5.3 mm (referenced to Bregma), ML:  $\pm$ 0.7 mm (referenced to the midline), and DV: -7.0 mm and -8.2 mm for male and -6.7 mm and -7.9 mm for female (referenced to the brain surface). Virus was obtained from the University of North Carolina at Chapel Hill Gene Therapy Center, courtesy of Dr. Karl Deisseroth. A total of 1~1.5  $\mu$ l of virus with a titer of  $10^{12}$  vg/ml was injected at the rate of 0.1  $\mu$ l/min per injection site. The rats also were implanted with optic fibers bilaterally (200 $\mu$ m diameter, Thorlab, NJ; AP: -5.3 mm, ML:  $\pm$  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 photobeam. A food dispenser containing 45 mg sucrose pellets (plain, banana-flavored, or grape-flavored; Bio-serv, Frenchtown, 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 were 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. Inter-trial intervals (ITI's) were 150-s. For all conditioning, V consisted of a cue light, and A1, A2 and A3 consisted of a tone, 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 O1, and A2 terminated with delivery three pellets of O2. A3 was paired with no food. After completion of the 12 days of simple conditioning, rats received four consecutive days of compound conditioning in which A1 and V were presented together as a 30-s compound cue terminating with six pellets of O1, and V, A2, and A3 continued to be presented as in simple conditioning. Cues were again presented in a blocked design, with order counterbalanced. For each cue, there were 12 trials on the first three days of compound conditioning and six trials on the last day of compound conditioning. During these sessions, light (532 nm, 16–18 mW output, Shanghai Laser & Optics Century Co., Ltd) was delivered into VTA during the reward period after presentation of the A1/V compound cue. To approximate the presumed normal timing of negative prediction errors in the Pavlovian over-expectation task as run in our lab, we made the assumption that negative prediction errors would be signaled not when reward was or was not delivered but rather when the rats would normally have encountered the additional pellets. Thus we delivered light pulses starting 6 s after reward delivery, based on the observation that rats spent at least 6 s (average  $7.3 \pm 1.3$  s,  $n = 6$  rats across 4 independent trials in pilot testing) in the food cup to consume 3 pellets versus 15 s ( $14.8 \pm 1.2$  s,  $n = 6$  rats across 3 independent test trials in pilot testing) to consume 6 pellets. Thus we reasoned that in the normal task, the rats would notice the absence of the additional reward during the period from around 6 to 15s and delivered our light pulses (three 2-s pulses, separated by 1.5-s) in this 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.

One day after the last compound conditioning session, rats received a probe test session consisting of eight non-reinforced presentations of A1, A2 and A3 stimuli, with the order mixed and counterbalanced. Subsequently all rats received 2 days of reminder training, which were identical to initial conditioning, and then the compound training and probe test were repeated, but with light delivered at the mid-point of the ITI (75 s after the end of the food pellet delivery period) following A1/V trials rather than during the critical reward period.

### Response measures

The primary measure of conditioning to cues was the percentage of time that each rat spent with its head in the food cup during the 30-s CS presentation, as indicated by disruption of the photocell beam. We also measured the percentage of time that each rat showed rearing behavior during the 30-s CS period. To correct for time spent rearing, the percentage of responding during the 30-s CS was calculated as: % of responding =  $100 * ((\% \text{ of time in food cup}) / (100 - (\% \text{ of time of rearing})))$ . Rearing was corrected for in order to reduce variability; there were no significant effects of group or experimental run on rearing behavior ( $F$ 's  $< 0.25$ ,  $p$ 's  $> 0.92$ . Fig S3). We also measured the amount of time the rats spent in the food cup during the 30-s period after CS presentation, during which food was

delivered and consumed, as a control for any aversive or distracting effect of optogenetic stimulation.

### Histology and Immunohistochemistry

Rats that received viral infusions and fiber implants were euthanized with an overdose of isoflurane and perfused with  $1 \times$  Phosphate buffered saline (PBS) followed by 4 % Paraformaldehyde (Santa Cruz Biotechnology Inc., CA). Fixed brains were cut in 40  $\mu\text{m}$  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 X-100/1  $\times$  PBS and then incubated in anti-tyrosine hydroxylase (TH) antisera (MAB318, 1: 600, EMD Millipore, Billerica, Massachusetts) followed by Alexa 568 secondary antisera (A11031, 1:1000, 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 pigmental area) and posterior (caudal, parabrachial pigmental 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  $\mu\text{m}$   $\times$  40  $\mu\text{m}$  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  $\mu\text{m}$ , co-localized within cells reactive to DAPI staining.

### Ex vivo electrophysiology

Three additional Th-Cre rats that received AAV-Ef1 $\alpha$ -DIO-NpHR3.0-eYFP injections in the VTA were anesthetized with isoflurane and perfused transcardially with  $\sim 40$  ml ice-cold NMDG-based artificial CSF (aCSF) solution containing (in mM): 92 NMDG, 20 HEPES, 2.5 KCl, 1.2 NaH<sub>2</sub>PO<sub>4</sub>, 10 MgSO<sub>4</sub>, 0.5 CaCl<sub>2</sub>, 30 NaHCO<sub>3</sub>, 25 glucose, 2 thiourea, 5 Na-ascorbate, 3 Na-pyruvate, and 12 N-acetyl-L-cysteine (300~310 mOsm, pH 7.3~7.4). After perfusion, the brains were immediately removed and the horizontal brain slices containing the VTA were made in 300  $\mu\text{m}$  with VT-1200 vibratome (Leica, Nussloch, Germany). The brain slices were recovered within 12 min at 32  $^{\circ}\text{C}$  in NMDG-based aCSF and incubated for at least 1 hr in HEPES-based aCSF containing (in mM): 92 NaCl, 20 HEPES, 2.5 KCl, 1.2 NaH<sub>2</sub>PO<sub>4</sub>, 1 MgSO<sub>4</sub>, 2 CaCl<sub>2</sub>, 30 NaHCO<sub>3</sub>, 25 glucose, 2 thiourea, 5 Na-ascorbate, 3 Na-pyruvate, and 12 N-acetyl-L-cysteine (300~310 mOsm, 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 NaH<sub>2</sub>PO<sub>4</sub>, 1 MgCl<sub>2</sub>, 2.4 CaCl<sub>2</sub>, 26 NaHCO<sub>3</sub>, 11 glucose, 0.1 picrotoxin, and 2 kynurenic acid, and saturated with 95% O<sub>2</sub>, and 5% CO<sub>2</sub> at 32 $^{\circ}\text{C}$ ~34 $^{\circ}\text{C}$ . Glass pipette (pipette resistance 2.0~2.5 M $\Omega$ , King Precision Glass, Claremont, CA) filled with K<sup>+</sup> based internal solution (in mM: 140 KMeSO<sub>4</sub>, 5 KCl, 0.05 EGTA, 2 MgCl<sub>2</sub>, 2 Na<sub>2</sub>ATP, 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<sup>+</sup> 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 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-second 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 posthoc immunohistochemical verification of TH expression. Throughout the recording, series resistance (10–15 M $\Omega$ ) 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 20k Hz and filtered at 10k Hz. Data were acquired in Clampex 10.3 (Molecular Devices, Foster City, CA), and was 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, 1:1500, 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.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This work was supported by the Intramural Research Program at the National Institute on Drug Abuse. The authors would like to thank Dr Karl Deisseroth and the Gene Therapy Center at the University of North Carolina at Chapel Hill for providing viral reagents, and Dr Garret Stuber for technical advice on their use. We would also like to thank Dr Brandon Harvey and the NIDA Optogenetic and Transgenic Core and Dr Marisela Morales and the NIDA Histology Core for their assistance. The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

## References

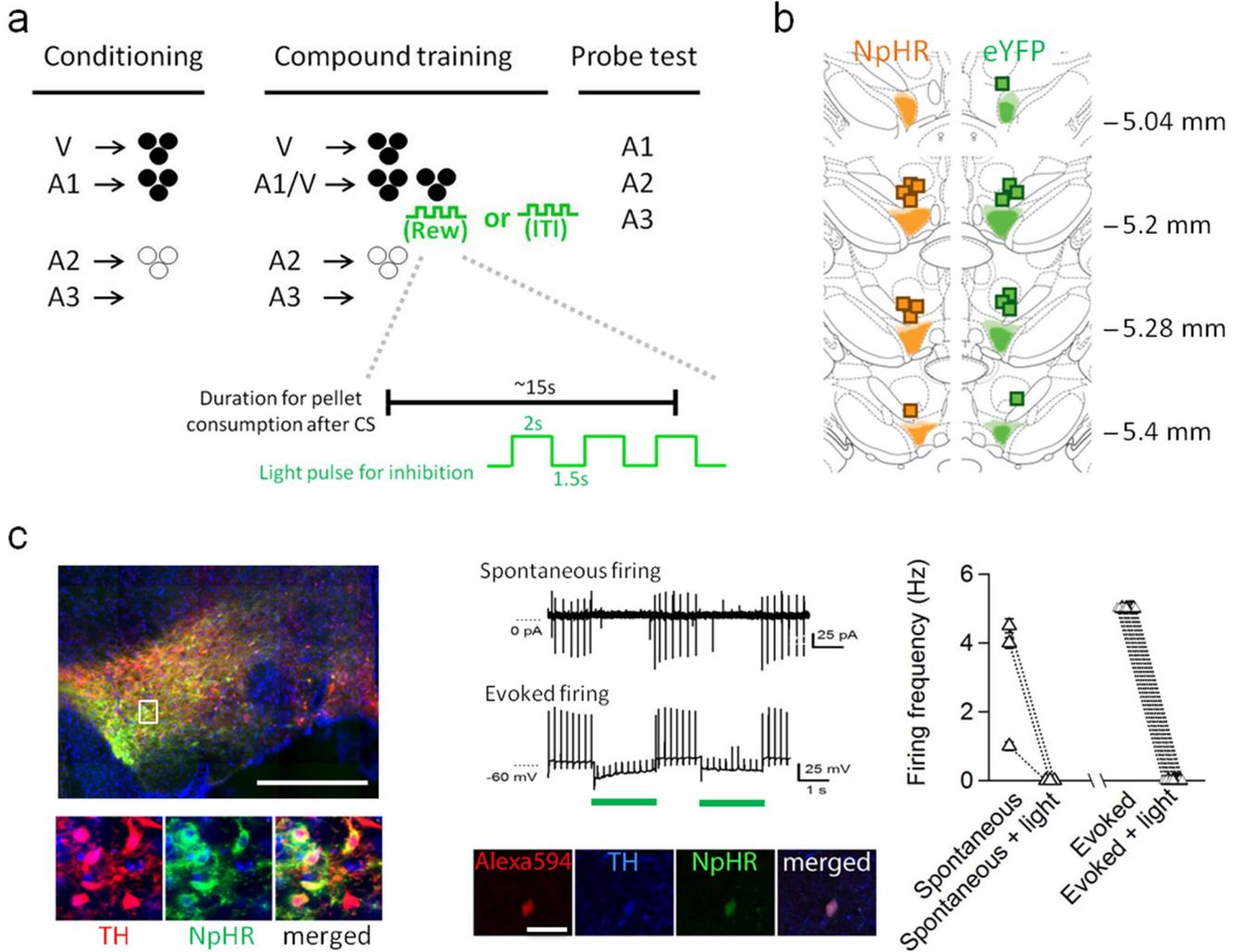
1. Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. *Classical Conditioning II: Current Research and Theory*. Appleton-Century-Crofts; New York: 1972. p. 64-99.
2. Sutton RS. Learning to predict by the method of temporal difference. *Machine Learning*. 1988; 3:9–44.
3. Mirenowicz J, Schultz W. Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology*. 1994; 72:1024–1027. [PubMed: 7983508]
4. Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nature Neuroscience*. 2007; 10:1615–1624. [PubMed: 18026098]
5. Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature*. 2001; 412:43–48. [PubMed: 11452299]
6. Pan WX, Schmidt R, Wickens JR, Hyland BI. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *Journal of Neuroscience*. 2005; 25:6235–6242. [PubMed: 15987953]

7. D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*. 2008; 319:1264–1267. [PubMed: 18309087]
8. Day JJ, Roitman MF, Wightman RM, Carelli RM. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nature Neuroscience*. 2007; 10:1020–1028. [PubMed: 17603481]
9. Hart AS, Rutledge RB, Glimcher PW, Phillips PE. Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. *Journal of Neuroscience*. 2014; 34:698–704. [PubMed: 24431428]
10. Schultz W, Dayan P, Montague PR. A neural substrate for prediction and reward. *Science*. 1997; 275:1593–1599. [PubMed: 9054347]
11. Redgrave P, Gurney K, Reynolds J. What is reinforced by phasic dopamine signals? *Brain Research Reviews*. 2008; 58:322–339. [PubMed: 18055018]
12. Zweifel LS, et al. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proceedings of the National Academy of Science*. 2009; 106:7281–7288.
13. Steinberg EE, et al. A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*. 2013; 16:966–973. [PubMed: 23708143]
14. Tsai HC, et al. Phasic firing in dopamine neurons is sufficient for behavioral conditioning. *Science*. 2009; 324:1080–1084. [PubMed: 19389999]
15. Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchinson KE. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Science*. 2007; 104:16311–16316.
16. Kim KM, et al. Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. *PLoS One*. 2012; 7:e33612. [PubMed: 22506004]
17. Stopper CM, Tse MT, Montes DR, Wiedman CR, Floresco SB. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron*. 2014; 84:177–189. [PubMed: 25220811]
18. Shumake J, Ilango A, Scheich H, Wetzel W, Ohl FW. Differential neuromodulation of acquisition and retrieval of avoidance learning by the lateral habenula and ventral tegmental area. *Journal of Neuroscience*. 2010; 30:5876–5883. [PubMed: 20427648]
19. Stamatakis A, Stuber GD. Activation of lateral habenula inputs to the ventral midbrain promotes behavioral avoidance. *Nature Neuroscience*. 2012; 15:1105–1107. [PubMed: 22729176]
20. Danna CL, Shepard PD, Elmer GI. The habenula governs the attribution of incentive salience to reward predictive cues. *Frontiers in Human Neuroscience*. 2013; 9:781. [PubMed: 24368898]
21. Bayer HM, Lau B, Glimcher PW. Statistics of midbrain dopamine neuron spike trains in the awake primate. *Journal of Neurophysiology*. 2007; 98:1428–1439. [PubMed: 17615124]
22. Glimcher PW. Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Science*. 2011; 108:15647–15654.
23. Kakade S, Dayan P. Dopamine: generalization and bonuses. *Neural Networks*. 2002; 15:549–559. [PubMed: 12371511]
24. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 2009; 459:837–841. [PubMed: 19448610]
25. Pearce, JM.; Kaye, H.; Hall, G. Predictive accuracy and stimulus associability: development of a model for Pavlovian learning. In: Commons, ML.; Herrnstein, RJ.; Wagner, AR., editors. *Quantitative Analyses of Behavior*. Ballinger; Cambridge, MA: 1982. p. 241-255.
26. Esber GR, Haselgrove M. Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society B*. 2011; 278:2553–2561. [PubMed: 21653585]
27. Rescorla RA. Renewal from overexpectation. *Learning and Behavior*. 2007; 35:19–26. [PubMed: 17557388]
28. Rescorla RA. Spontaneous recovery from overexpectation. *Learning and Behavior*. 2006; 34:13–20. [PubMed: 16786880]

29. Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*. 2007; 191:507–520. [PubMed: 17031711]
30. Salamone JD, Correa M. The mysterious motivational functions of mesolimbic dopamine. *Neuron*. 2012; 76:470–485. [PubMed: 23141060]
31. Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*. 1998; 28:309–369. [PubMed: 9858756]
32. Ikemoto S, Yang C, Tan A. Basal ganglia circuit loops, dopamine and motivation: A review and enquiry. *Behavioral Brain Research*. 2015 pii: S0166–4328(15)00260–0. Epub ahead of print.
33. Takahashi Y, et al. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*. 2009; 62:269–280. [PubMed: 19409271]
34. Takahashi YK, et al. Neural estimates of imagined outcomes in the orbitofrontal cortex drive behavior and learning. *Neuron*. 2013; 80:507–518. [PubMed: 24139047]
35. Rescorla RA. Reduction in the effectiveness of reinforcement after prior excitatory conditioning. *Learning and Motivation*. 1970; 1:372–381.
36. Witten IB, et al. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*. 2011; 72:721–733. [PubMed: 22153370]
37. Lammel S, et al. Diversity of transgenic mouse models for selective targeting of midbrain dopamine neurons. *Neuron*. 2015; 85:429–438. [PubMed: 25611513]
38. Stamatakis AM, et al. A unique population of ventral tegmental area neurons inhibits the lateral habenula to promote reward. *Neuron*. 2013; 80:1039–1053. [PubMed: 24267654]
39. Stuber GD, Stamatakis AM, Kantak PA. Considerations when using cre-driver rodent lines for studying ventral tegmental area circuitry. *Neuron*. 2015; 85:439–445. [PubMed: 25611514]
40. Yamaguchi T, Qi J, Wang HL, Zhang S, Morales M. Glutamatergic and dopaminergic neurons in the mouse ventral tegmental area. *European Journal of Neuroscience*. 2015; 41:760–772. [PubMed: 25572002]
41. Li X, Qi J, Yamaguchi T, Wang HL, Morales M. Heterogenous composition of dopamine neurons of the rat A10 region: molecular evidence for diverse signaling properties. *Brain Structure and Function*. 2013; 218:1159–1176. [PubMed: 22926514]
42. Root DH, et al. Norepinephrine activates dopamine D4 receptors in the rat lateral habenula. *Journal of Neuroscience*. 2015; 35:3460–3469. [PubMed: 25716845]
43. Tecuapetla F, et al. Glutamatergic signaling by mesolimbic dopamine neurons in the nucleus accumbens. *Journal of Neuroscience*. 2010; 30:7105–7110. [PubMed: 20484653]
44. Stuber GD, Hnasko TS, Britt JP, Edwards RH, Bonci A. Dopaminergic terminals in the nucleus accumbens but not the dorsal striatum corelease glutamate. *Journal of Neuroscience*. 2010; 30:8229–8233. [PubMed: 20554874]
45. Zhang S, et al. Dopaminergic and glutamatergic microdomains in a subset of rodent mesoaccumbens axons. *Nature Neuroscience*. 2015; 18:386–392. [PubMed: 25664911]
46. Matsumoto M, Hikosaka O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*. 2007; 447:1111–1115. [PubMed: 17522629]
47. Hong S, Zhou TC, Smith M, Saleem KS, Hikosaka O. Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *Journal of Neuroscience*. 2011; 31:11457–11471. [PubMed: 21832176]
48. Ji H, Shepard PD. Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA(A) receptor-mediated mechanism. *Journal of Neuroscience*. 2007; 27:6923–6930. [PubMed: 17596440]
49. Mileykovskiy B, Morales M. Duration of inhibition of ventral tegmental area dopamine neurons encodes a level of conditioned fear. *Journal of Neuroscience*. 2011; 31:7471–7476. [PubMed: 21593330]
50. Ilango S, et al. Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *Journal of Neuroscience*. 2014; 34:817–822. [PubMed: 24431440]

### Summary

Phasic changes in dopamine activity correlate with prediction error signaling. But causal evidence that these brief changes in firing actually drive associative learning is rare. Here the authors show that brief pauses in dopamine neuron firing at the time of reward mimic the effects of endogenous negative prediction errors.



**Figure 1. Task design, fiber placements, and immunohistochemical and electrophysiological verification of Cre-dependent NpHR and eYFP expression in 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 spread of expression at each level, whereas the dark shading represents the minimal spread. c) Images (left) show that majority of NpHR-expressing neurons (green) also expressed tyrosine hydroxylase (red). Scale: 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 will inevitably render some area seems to be overpowered by the signal and tip 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 bar). These neurons also expressed TH, confirmed by intracellular labeling and post-hoc TH staining (middle bottom). Spontaneous and evoked

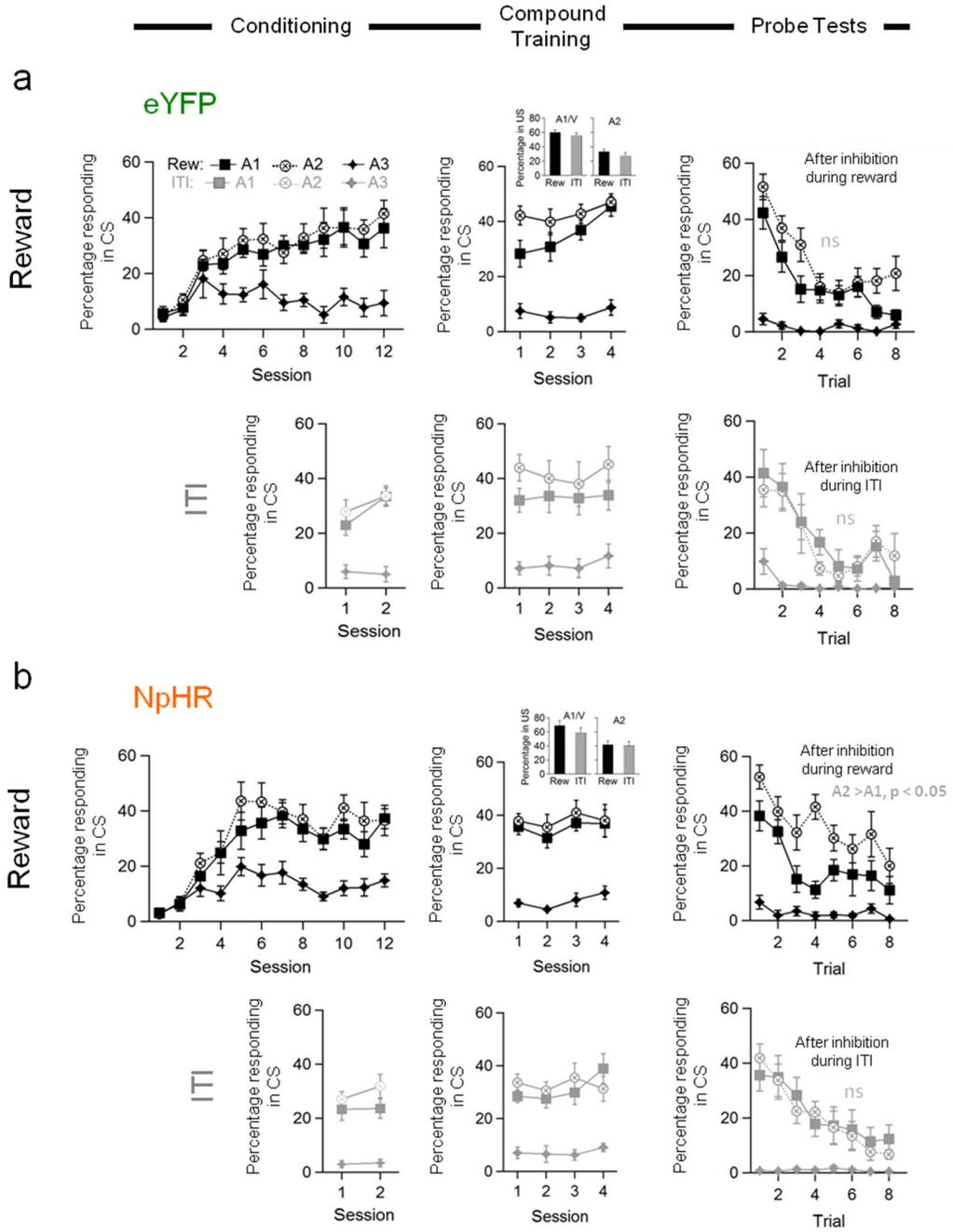
firing of NpHR-expressing neurons were interrupted by brief pulses of light inhibition (n = 10 from 3 subjects, firing activity of individual neurons is summarized at the right).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 2. Optogenetic inhibition of TH+ neurons in VTA mimics learning induced by reward over-expectation**

Conditioned responding of a) eYFP and b) NpHR rats to the critical auditory cues during the initial conditioning or reminder training (left column), during compound training (middle column), and the subsequent probe tests (right column). Conditioned responding is represented as the percent of time the rats spent in the food cup during the cues. Percent of time 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 (1<sup>st</sup> and 3<sup>rd</sup> rows, Reward), when TH+ neurons were inhibited at the time of reward, is shown in

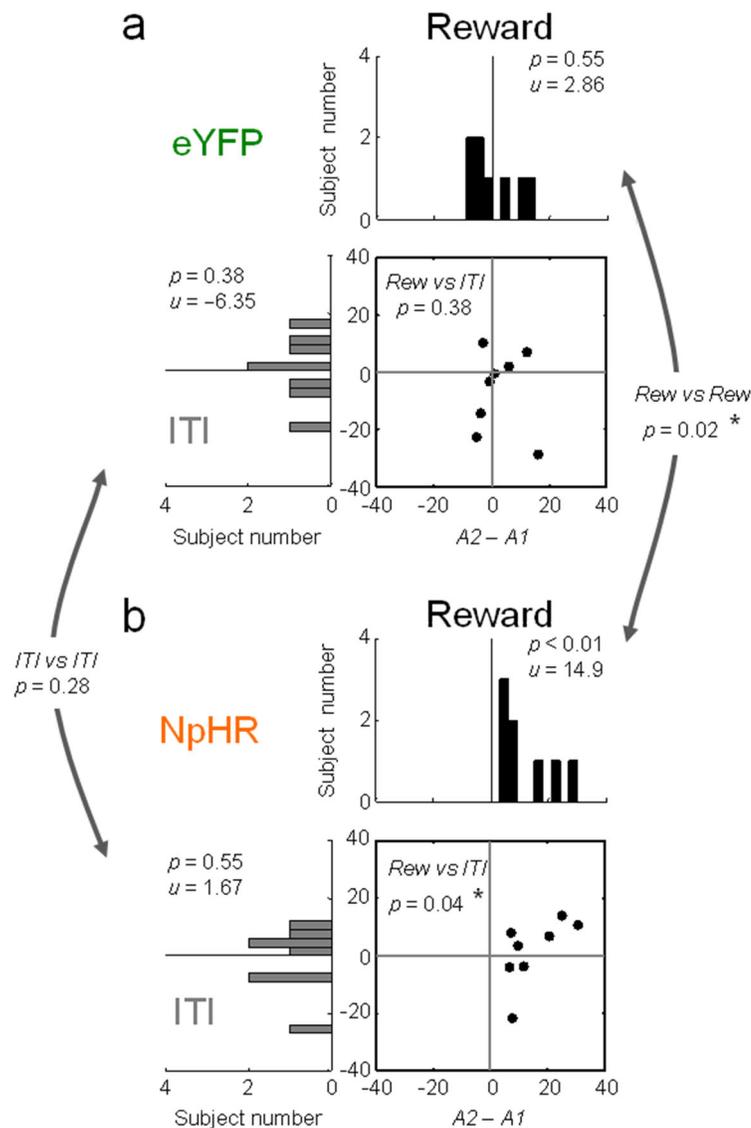
the dark symbols; data from the ITI run (2<sup>nd</sup> and 4<sup>th</sup> rows, ITI), when TH+ neurons were inhibited in the ITI is shown in the light symbols. N = 8 each group. Vertical bars show S.E.M. n.s., non-significant at > 0.10; \*, p < 0.0001 based on stats given in main text. All values in each line plot represent % time rats spent in food cup during the CS after correction for rearing.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 3. Differences in conditioned responding caused by optogenetic inhibition of TH+ neurons in VTA**

Differences in conditioned responding of a) eYFP and b) NpHR rats to the critical auditory cues in probe tests after inhibition during reward (black histograms) or after inhibition during the ITI (gray histograms). Scatter plots show data points from the individual rats in each group across the two tests. Difference scores were calculated for each rat as the difference in responding to A2-A1 across all trials in the relevant probe test (using data shown in Fig. 2). Distributions are centered on the difference on the final day of conditioning prior to compound training. The numbers in each panel indicate results of Wilcoxon signed-rank test ( $p$ ) and the average scores ( $u$ ). Numbers comparing panels indicate the results of a Wilcoxon rank sum test ( $p$ ).