



Published in final edited form as:

Nat Neurosci. 2017 May ; 20(5): 735–742. doi:10.1038/nn.4538.

Dopamine transients are sufficient and necessary for acquisition of model-based associations

Melissa J Sharpe^{1,2}, Chun Yun Chang¹, Melissa A. Liu¹, Hannah M. Batchelor¹, Lauren E. Mueller¹, Joshua L Jones¹, Yael Niv², and Geoffrey Schoenbaum^{1,3,4}

¹NIDA Intramural Research Program, Baltimore, MD 21224

²Department of Psychology and Neuroscience Institute, Princeton University, Princeton, NJ 08544

³Departments of Anatomy & Neurobiology and Psychiatry, University of Maryland School of Medicine, Baltimore, MD 21201

⁴Solomon H. Snyder Department of Neuroscience, The Johns Hopkins University, Baltimore, MD 21287

Abstract

Associative learning is driven by prediction errors. Dopamine transients correlate with these errors, which current interpretations limit to endowing cues with a scalar quantity reflecting the value of future rewards. Here, we tested whether dopamine might act more broadly to support learning of an associative model of the environment. Using sensory preconditioning, we show that prediction errors underlying stimulus-stimulus learning can be blocked behaviorally and reinstated by optogenetically activating dopamine neurons. We further show that suppressing the firing of these neurons across transition prevents normal stimulus-stimulus learning. These results establish that the acquisition of model-based information about transitions between non-rewarding events is also driven by prediction errors, and that contrary to existing canon, dopamine transients are both sufficient and necessary to support this type of learning. Our findings open new possibilities for how these biological signals might support associative learning in the mammalian brain in these and other contexts.

Introduction

The discovery that midbrain dopamine neurons emit a teaching signal when an unexpected reward or reward-predicting cue occurs has transformed how we conceptualize dopamine

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

Correspondence or requests for material should be addressed to MJS (melissa.sharpe@nih.gov) or GS (geoffrey.schoenbaum@nih.gov).

Supplementary Information is available on the web.

Author Contributions: MJS and GS designed the experiments, MJS, MAL, HMB, and LEM collected the data with technical advice and assistance from CYC and JLJ. MJS and GS analyzed the data. MJS, YN, and GS interpreted the data and wrote the manuscript with input from the other authors.

Competing Financial Interests Statement: Data available upon request. The authors have no competing interests to report.

function¹. The response to unpredicted rewards, initially large, wanes as the subject comes to anticipate the rewarding event, transferring instead to antecedent stimuli that reliably predict future reward. This finding has been influential because transient changes in dopamine are so like the prediction-errors argued to drive learning in reinforcement learning models²⁻⁵. Indeed, the dopaminergic prediction error has become almost synonymous with the reward prediction error defined in these models. However, these errors are thought to support only a relatively limited form of learning in which predictive cues are endowed with a scalar quantity that reflects the rewarding *value* of future events at the time of learning. This cached or model-free value does not capture any specific information about the identity of those future events, even in more expansive recent proposals that incorporate elements of reward structure⁴. As a result, the behaviors supported by these values are relatively inflexible, since they cannot reflect information about the predicted events other than their general value at the time of learning.

Yet much behavior reflects specific information about predicted events, rewarding or otherwise⁶. Such behavior reveals the existence of a rich and navigable associative representation or model of the structure of the environment. For instance, when walking into your favorite neighborhood restaurant, you expect not only a good meal, but also one that consists of sushi, not pasta. Because this prediction contains specific information beyond value, it supports flexible and adaptive behavior⁷⁻¹¹. You might love Japanese and Italian food equally, but if you become pregnant and are instructed to avoid raw fish, you can adjust your choice of restaurant without additional direct experience. Can dopaminergic prediction errors support the formation of these model-based associations or do they only support learning of model-free associations that contain scalar values? Although optogenetic studies have confirmed that dopamine transients can function as errors to support associative learning¹²⁻¹⁸, this critical question remains unaddressed, since in each of these experiments the resultant behavior could be accounted for by model-free learning mechanisms.

Here we directly address this question using sensory-preconditioning¹⁹⁻²². Sensory preconditioning entails presenting subjects with two neutral cues, e.g., C and X, in close succession, such that a predictive relationship $C \rightarrow X$ can form between them. Importantly, in this preconditioning phase, no rewards are delivered, and consequently no new behavioral responses or scalar values are learned. However, the contents of what is learned in preconditioning can be revealed if the second cue is subsequently paired with an unconditional stimulus, for instance, a reward (i.e. $X \rightarrow US$). Subsequently, both C and X will elicit robust conditioned responses. Since C was never paired with reward, the response to C demonstrates the existence of an associative link between C and X. The use of this $C \rightarrow X$ association to support responding for the reward is a classic example of model-based behavior.

We used this behavioral approach in two experiments. The first was designed to test whether a dopamine transient is sufficient to support the formation of the associative representations underlying model-based behavior. For this, we combined sensory preconditioning with blocking²³, a procedure developed to show that associative learning depends on the presence of a prediction error. While blocking has previously been shown only in the context of learning about a valuable reward²³, we hypothesized that learning of associations that do

not involve reward or value should also be regulated by an error mechanism. To test this, we applied the same logic used in reward blocking to reduce acquisition of the $C \rightarrow X$ relationship during preconditioning. In particular, we first paired a different cue, A, with X ($A \rightarrow X$). Then, during preconditioning, A was presented in compound with C, followed by X ($AC \rightarrow X$). Because A already predicts X, if learning of the stimulus-stimulus association is driven by errors in prediction, the presence of A should diminish or block the formation of any association between C and X. Indeed, we observed such blocking in pilot testing (see Supplementary Information, Fig S1), confirming that the initial learning in sensory preconditioning is driven by prediction errors (termed state prediction errors in current computational models⁸), even though there is no reward or value present.

Against this background, we attempted to reinstate learning of the $C \rightarrow X$ association by briefly activating the dopamine neurons at the start of the X cue in the $AC \rightarrow X$ trials, using parameters designed to evoke firing similar to that sometimes observed for rewards^{13, 24–28} or even neutral cues^{24, 29, 30}. We reasoned that if dopamine transients can support learning of associations between the neural representations of events in the environment, as opposed to being restricted to the addition or subtraction of value, then this manipulation should restore normal sensory preconditioning of C. In a second experiment, we tested for necessity of dopamine for this learning process by suppressing the dopamine neurons across the transition between the cues during a standard sensory preconditioning task. The results of the two experiments show that dopamine transients are both sufficient and likely necessary to support the acquisition of the associative structures underlying model-based behavior.

Results

Dopamine transients are sufficient for the formation of model-based associations

Prior to training, all rats underwent surgery to infuse virus and implant fiber optics targeting the ventral tegmental area (VTA; Fig 1). We infused AAV5-EF1 α -DIO-ChR2-eYFP (ChR2 experimental group; $n=18$) or AAV5-EF1 α -DIO-eYFP (eYFP control group; $n=19$) into the VTA of rats expressing Cre recombinase under the control of the tyrosine hydroxylase (TH) promoter³¹. After surgery and recovery, rats were food restricted until their body weight reached 85% of baseline and training commenced. Training then began with two days of preconditioning. On the first day, the rats received a total of 16 pairings of two 10s neutral cues ($A \rightarrow X$). On the second day, the rats continued to receive pairings of the same two neutral cues ($A \rightarrow X$; 8 trials). In addition, on other trials, the first cue was presented together with a second, novel neutral cue (either $AC \rightarrow X$ or $AD \rightarrow X$; 8 trials each). On AC trials, blue light (473nm, 20Hz, 16–18mW output, Shanghai Laser & Optics Century Co., Ltd) was delivered for 2s at the start of X to activate VTA dopamine neurons. As a temporal control for non-specific effects, the same light pattern was delivered on AD trials in the inter-trial interval, 120–180s after termination of X. Finally, to verify that sensory preconditioning could be obtained with compound cues, the rats also received pairings of two novel 10s cues with X ($EF \rightarrow X$; 8 trials). As expected, since training did not involve pairing with reward, rats in both groups (ChR2 and eYFP controls) exhibited little responding at the food cup during any of the cues on either day of training (Fig 2a); a two-factor ANOVA on food cup

entries during cue presentations (cue \times group) revealed no main effect ($F(4,140) = 1.52, p = 0.2$), nor any interaction with group ($F(4,140) = 0.276, p = 0.893$).

Following preconditioning, the rats began conditioning, which continued for 4 days. Each day, the rats received 24 trials in which X was presented followed by delivery of two 45mg sucrose pellets (X \rightarrow 2US). Rats in both groups acquired a conditioned response to X. This was evident as an increase in the number of times they entered the food cup to look for sucrose pellets during X, across days of conditioning (Fig 2b). Importantly, the acquisition of this conditioned response was similar in the two groups; a two-factor ANOVA (group \times day) revealed a main effect of day ($F(3,105) = 39.71, p < 0.0001$) but neither main effect ($F(1,35) = 0.553, p = 0.46$) nor any interaction with group ($F(3,105) = 0.13, p = 0.94$). Thus, the introduction of a dopamine transient at the start of X did not produce any lasting effect on subsequent processing of or learning about X.

Finally, the rats received a probe test in which each of the critical test cues (C, D, F) were presented 4 times each, in an interleaved and counterbalanced order, alone and without reward. This probe test was designed to assess whether these preconditioned cues had acquired the ability to predict sucrose pellet delivery. As expected from studies of normal sensory preconditioning, rats in both groups demonstrated high levels of responding to F, suggesting that, despite the use of a compound cue, they learned that F predicted X and used that relationship in the probe test to infer that F predicted sucrose pellets (Fig 2c, black bars). Rats in both ChR2 and eYFP groups also demonstrated low levels of responding to D (as in our pilot study, see Supplementary Information, Fig S1), indicating that the presence of A and its ability to predict X had blocked D from becoming associated with X (Fig 2c, red bars). Notably, this occurred despite transient activation of the VTA dopamine neurons during the inter-trial interval following AD trials. A two-factor ANOVA (cue \times group) on responding during presentation of cues F and D revealed a main effect of cue ($F(1,35) = 4.372, p = 0.044$) but no main effect ($F(1,35) = 0.001, p = 0.982$) or interaction with group ($F(1,35) = 0.287, p = 0.595$). Thus, both groups exhibited identical blocking of sensory preconditioning as indexed by a significant difference between F and D.

When delivered at the start of X on the AC trials, however, the transient activation of the dopamine neurons unblocked learning, so that responding to C was higher than responding to D in the ChR2 group but not in the eYFP controls (Fig 2c, blue bars). A two-factor ANOVA (cue \times group) on responding to C and D revealed a main effect of cue ($F(1,35) = 4.599, p = 0.039$) and a significant interaction with group ($F(1,35) = 4.154, p = 0.049$). This interaction was due to a significant difference between responding to C and D in the ChR2 group ($F(1,35) = 8.52, p = 0.006$) but not in the eYFP group ($F(1,35) = 0.006, p = 0.940$). In addition, responding to D did not differ between groups ($F(1,35) = 0.153, p = 0.698$), whereas responding to C was significantly higher in the ChR2 rats than in the eYFP controls ($F(1,35) = 5.277, p = 0.028$). Thus, transient activation of the VTA dopamine neurons at the start of X on AC trials reversed the blocking effect, as indexed by the significant increase in responding to C only in the ChR2 rats.

But is the learning supported by transient activation of dopamine neurons the same as what is normally learned during sensory preconditioning? That is, did the rats in the ChR2 group

respond to C because it evoked a prediction that sucrose pellets would be delivered to the food cup? To test this, we assessed the effect of devaluation of the sucrose pellets on responding to C in a subset of the ChR2 rats that had been trained on the blocking of sensory preconditioning task. We divided the rats into two groups with equal responding to C ($F_{(1,8)} = 0.028$, $p=0.871$). After reminder training ($X \rightarrow 2US$; 12 trials; $F_{(1,8)} = 2.802$, $p = 0.133$), rats in each group received sucrose pellets and lithium chloride injections to induce nausea (LiCl; 10ml/kg 0.15M) on three successive days. For one group (devalued group; $n = 5$), sucrose pellets were presented immediately prior to induction of illness. For the other group (non-devalued group; $n = 5$), sucrose pellets were presented ~6h after the induction of illness. Two days after the final LiCl injection, the rats received a probe test in which C was presented as before, alone and without reward. In this test, rats in the devalued group responded significantly less to C than rats in the non-devalued group (12 trials; Fig 3a; $F_{(1,8)} = 6.777$, $p = 0.031$). Devalued rats also consumed fewer sucrose pellets during a subsequent consumption test (Fig 3b; $F_{(1,8)} = 13.425$, $p = 0.006$), confirming a reduced desire for the pellets. The effect of devaluation on responding to C in the ChR2 rats is the same as what has been previously reported for a normally preconditioned cue^{20, 22}, suggesting that activating dopamine neurons transiently at the start of X on the AC trials restored normal acquisition of the predictive relationship between C and X, effectively leading to anticipation of sucrose pellets upon presentation of C.

Dopamine transients are necessary for the formation of model-based associations

The above shows that transient activation of VTA dopamine neurons is sufficient to drive the formation of an association between two sensory representations. This association can then support model-based behavior, with rats responding to C as if it predicts food through its association with X. This is important because we know that dopamine neurons exhibit transient increases in firing in the context of unexpected reward. The results described above suggest the dopamine transient at the time of an unexpected reward should result in an association between the cue and the sensory features of the reward that could later be used to support devaluation-sensitive behavior or even economic decision-making.

Of course, the finding above does not address whether transient activation of these neurons normally contributes to sensory preconditioning or stimulus-stimulus learning in the absence of reward. Although the timing and duration of the optogenetic activation was designed based on the dopamine responses to reward^{13, 24–28, 32–34}, its duration is longer than the peak response typically observed in unit studies. Further, while dopamine neurons have been shown to fire in response to neutral cues^{24, 29, 30}, such activity is weaker than to unexpected rewards. Therefore it is not clear how similar the signal that our stimulation generated was to that caused by unexpected sensory input in the absence of reward. Further, idiosyncrasies governing viral expression and light penetration dictate that no pattern of optogenetic activation is likely to reproduce what happens normally, either here or in other similar work.

To address whether dopamine transients are necessary for model-based learning in the absence of reward, we ontogenetically suppressed activity in VTA dopamine neurons across the critical transition between the sensory cues in the first phase of a standard sensory preconditioning task. Rats were presented with two pairs of neutral cues in close succession

(i.e. $A \rightarrow X$; $B \rightarrow Y$). Dopamine neurons were prevented from firing during the transition between B and Y, but were free to fire between A and X. Subsequently, X and Y were paired directly with reward ($X \rightarrow US$; $Y \rightarrow US$). We reasoned that if dopamine transients were necessary for learning associations between non-rewarding events in the environment, then suppressing the firing of dopamine neurons across this transition would disrupt normal sensory preconditioning of B.

Prior to training, all rats underwent surgery to infuse virus and implant fiber optics targeting the VTA (Fig 4). We infused AAV5-EF1 α -DIO-eNpHR3.0-eYFP (NpHR experimental group; $n=17$) or AAV5-EF1 α -DIO-eYFP (eYFP control group; $n=24$) into the VTA of rats expressing Cre recombinase under the control of the TH promoter³¹. Note that because of the much higher amount of reward in this experiment versus the first experiment (approximately double), the nature of the conditioned response was different in this experiment. Rather than checking briefly many times for reward, the rats made fewer entries and spent more time inside the food cup. As a result, although we observed similar effects on both measures, here we plotted conditioned responding as the amount of time spent in the food cup rather than number of entries (see comment on response measures in Online Methods and Supplementary Information, Fig S2 for more details).

After surgery and recovery, rats were food restricted until their body weight reached 85% of baseline, and training commenced. Training began with a day of preconditioning. Rats received a total of 12 pairings of two 10s neutral cues ($B \rightarrow Y$). On $B \rightarrow Y$ trials, continuous green light (532nm, 16–18mW output, Shanghai Laser & Optics Century Co., Ltd) was delivered for 2.5s beginning 500ms before the termination of B continuing across the start of Y for 2s in order to inactivate VTA dopamine neurons across a time window that would prevent any transient increase in activity of these neurons at the beginning of X. As a positive control, the rats also received 12 pairings of two other novel 10s cues during this phase ($A \rightarrow X$; 12 trials). No light was delivered across $A \rightarrow X$ pairings. As no rewards were delivered during this phase of training, rats in both groups (eYFP and NpHR) exhibited very little responding at the food cup during cue presentation (Fig 5a); a two-factor ANOVA on food cup responding during cue presentations (cue \times group) revealed no main effect ($F(1,39) = 1.88$, $p = 0.177$) nor any interaction with group ($F(3, 117) = 0.425$, $p = 0.736$).

Following preconditioning, the rats began conditioning, which continued for 4 days. Each day, the rats received 24 trials in which X and Y were both presented followed by delivery of two 45mg sucrose pellets ($X \rightarrow 2US$). X was paired with one flavor of sucrose pellet, whereas Y was paired with another flavor (banana or grape, counterbalanced). Rats in both groups acquired a conditioned response to X and Y, as evident from the increase in the percentage of time they spent in the food cup during X and Y in expectation of sucrose pellets across days of conditioning (Fig 5b). The acquisition of this conditioned response was similar in the two groups and for both cues; a three-factor ANOVA (cue \times group \times day) revealed a main effect of day ($F(3,105) = 43.181$, $p < 0.0001$) but no main effect of cue ($F(1,39) = 0.008$, $p = 0.927$), or group ($F(1,39) = 0.094$, $p = 0.761$) or any cue \times group interaction ($F(1,39) = 1.113$, $p = 0.298$). Thus, suppression of dopaminergic activity across the transition between B and Y did not produce lasting effects on processing of or learning about Y.

Lastly, the rats received a probe test in which each of the critical test cues (A and B) were presented 6 times each, in an interleaved and counterbalanced order, alone and without reward. As expected, rats in the eYFP group exhibited equally high levels of conditioned responding to both A and B (Fig 5c, top panel), showing that regardless of light delivery, they learned the predictive relationship between both cue pairs and used them in the probe test to predict the delivery of sucrose pellets. By contrast, rats in the NpHR group exhibited significantly lower conditioned responding to B, the cue at the end of which we suppressed the dopamine neurons, than to A, the control cue. A two-factor ANOVA (cue \times group) revealed a main effect of cue ($F_{(1,39)}=5.94$, $p=0.019$) and a significant cue by group interaction ($F_{(1,39)}=4.68$, $p=0.037$). Subsequent comparisons showed that this interaction was due to a significant difference in responding to cues A and B in the NpHR group ($F_{(1,39)}=4.952$, $p=0.012$) that was not present in the eYFP group ($F_{(1,39)}=0.742$, $p=0.483$). Notably, this within-subject difference could not be explained by the slightly (but not significantly) increased responding to A in the NpHR group (see Supplementary Information, Fig S3 for additional information on high versus low responders). This difference is consistent with the proposal that by preventing transient activation of the dopamine neurons at the B \rightarrow Y transition, we prevented formation of the normal association between these two cues.

Discussion

We have shown that activity in VTA dopamine neurons is sufficient and necessary for the formation of associative structures that underlie model-based behavior. In our first experiment, we demonstrated that transient activation of dopaminergic neurons, with a timing and duration designed to mimic a prediction error, unblocked stimulus-stimulus learning in a sensory preconditioning task, resulting in later responding that reflected a prediction of sucrose pellet delivery that could not have been directly acquired under the influence of the artificial dopamine transient we induced. In the second experiment, we demonstrated that suppressing dopamine neurons, with a timing and duration designed to interfere with any dopamine transients, blocked stimulus-stimulus learning in a sensory preconditioning task.

Current conceptualizations of dopamine transients as the reward prediction errors postulated by model-free reinforcement learning algorithms cannot explain these data. This is because the error signal in these models functions only to endow the predictive cue with a scalar quantity that reflects the value of future events; the resultant associative representation does not incorporate or link to specific information about the identity of these events beyond their value at the time of learning. As a result, this type of learning cannot explain why the rats in the first experiment searched in the food cup for sucrose pellets when C was presented in the probe test, and then stopped doing so when the pellets were no longer desirable. Even if the dopamine transients endowed C with cached value (as reinforcement-learning models propose), and this was the reason for the food cup responding, such effects on behavior would generalize beyond the specific reward and therefore be insensitive to its devaluation³⁵. Indeed, if we stimulated dopamine to unblock learning when food was present, as has been done¹², these models predict that resultant responding would be insensitive to devaluation. Likewise responding to C in our experiment also cannot reflect

direct reinforcement of the motor response by the dopamine transient, since, in contrast to even the most well-controlled prior studies^{12, 14}, this response was not present when dopamine neuron activity was manipulated. Of course, such non-specific responding would also be insensitive to devaluation of the food reward³⁶, contrary to our results. Thus our results go far beyond what can be explained by a cached-value prediction error.

Nor can the results from either experiment reflect changes in salience or associability caused by manipulation of the dopamine neurons, either directly or via the addition or subtraction of cached value. While such effects have been reported following optogenetic activation of dopamine terminals in medial prefrontal cortex³⁷, we saw no evidence of this in either of our experiments involving manipulation of the cell bodies. For example, while increasing the salience or associability of X on the AC trials in our first experiment might indirectly allow X to enter into an association more readily with C, all theoretical accounts of which we are aware^{38–40} would also predict lasting effects on processing and associability of X. These effects would facilitate learning for X in other parts of our task, and yet we did not observe any evidence of increased learning about X in other trials in the ChR2 rats. In particular, the ChR2 rats did not respond more to D than controls, nor did they show more rapid conditioning to X in the second phase of training. The same is true for our second experiment, where we did not see any changes in learning about Y during conditioning, indicating that suppressing dopamine neurons did not alter the salience or value of Y. It is also worth noting that direct effects on salience would be inconsistent with evidence that activation of VTA dopamine neurons diminishes extinction learning while inhibition of these neurons facilitates it^{12, 14}. These effects, achieved using the same optogenetic approaches applied here, are the opposite of what would be expected if manipulating these neurons directly altered salience.

Instead, the most parsimonious explanation of our results is that dopamine transients play a role in the formation of associative links between the neural representations of external events, whether rewarding or not – linking representations of neutral cues during preconditioning and representations of neutral cues with representations of rewards in other settings. Importantly, this interpretation holds whether the ultimate behavior in the probe test reflects inference (i.e. if $A \rightarrow X$ and $X \rightarrow US$, then $A \rightarrow US$) or mediated learning during the conditioning phase (i.e. X evokes a memory of A that becomes directly associated with the US, so that later $A \rightarrow US$) (see Supplementary Information, Fig S1 caption for further consideration). In either case, dopamine must be acting to influence the association between the cues in the first phase of training. While this proposal does not negate a role for dopamine in learning about cached values, it does represent a substantial expansion of the kind of learning that dopaminergic prediction errors are thought to support. Along with recent data showing that these prediction-error signals can reflect value predictions derived from model-based associative structures^{11, 41–45}, our results show that dopaminergic error signals are potentially richer, more complex, and more capable than previously envisioned.

This is good news, given how difficult it has been to find plausible candidate neural substrates to signal these other types of prediction errors; the dopamine neurons appear relatively unique in the strength of their error signaling⁴⁶. Of course, our experimental approach affects a general population of VTA dopamine neurons that likely projects broadly

to multiple target regions. The neurons activated are determined somewhat at random, based on viral expression and light penetration. In this sense, our manipulations – both the activation as well as the suppression – are not, strictly speaking, physiological. This caveat is important to keep in mind when evaluating the importance of this or any other similar study. One way to view the ability of these manipulations to produce principled results is that the relatively simple and highly constrained behavioral designs allow us to see real effects despite our poor ability to truly reproduce real-world patterns of activity. We would speculate that in normal settings, the precise sort of associative information that is acquired under the influence of dopaminergic error signals will presumably reflect subtle variations in the content of the signal^{33, 47} combined with specialization of the downstream region or circuit^{21, 48, 49}.

Finally, it is worth noting that our results represent the first demonstration of which we are aware that learning about neutral cues is regulated by prediction errors. That is, in our blocking of sensory preconditioning procedure, we found that prior learning of the association between A and X blocked the ability of animals to learn that D predicts X. This shows that learning to associate neutral cues reflects contingency and not just contiguity between the two cues, matching previous demonstrations of blocking for cues predictive of motivationally-significant outcomes²³. That dopamine transients are both sufficient and necessary for this type of learning is in accord with observations that dopamine neurons exhibit error-like responses to novel or unexpected neutral cues under some conditions^{24, 29, 30}. Rather than reflecting a “novelty bonus”, such responses may reflect the informational prediction errors that are available in these circumstances to drive the sort of learning we have isolated here. Viewed from this perspective, the classic reward prediction errors normally observed in the firing of individual dopamine neurons might be a special (and especially strong) example of a more general function played by dopaminergic ensembles in signaling errors in event prediction. To determine if this is true, it will be necessary to interrogate dopaminergic activity in more complex behavioral paradigms, in which the source of the errors can be manipulated independent of value. In addition, it will likely be important to monitor groups of dopamine neurons in real time, using approaches such as calcium imaging³³, to identify information represented across neurons as has been done effectively to understand the functions of other brain regions⁵⁰.

Online Methods

Surgical Procedures

Rats received bilateral infusions of 1.2 μ L AAV5-EF1 α -DIO-ChR2-eYFP ($n=18$), AAV5-EF1 α -DIO, eNpHR3.0-eYFP ($n=17$), or AAV5-EF1 α -DIO-eYFP ($n=43$) into the VTA at the following coordinates relative to bregma: AP: -5.3 mm; ML: ± 0.7 mm; DV: -6.5 mm and -7.7 (females) or -7.0 mm and -8.2 mm (males). Virus was obtained from the Vector Core at University of North Carolina at Chapel Hill (UNC Vector Core). During surgery, optic fibers were implanted bilaterally (200 μ m diameter, Precision Fiber Products, CA) at the following coordinates relative to bregma: AP: -5.3 mm; ML: ± 2.61 mm, and DV: -7.05 mm (female) or -7.55 mm (male) at an angle of 15° pointed toward the midline.

Apparatus

Training was conducted in 8 standard behavioral chambers (Coulbourn Instruments; Allentown, PA), which were individually housed in light- and sound-attenuating boxes (Jim Garmon, JHU Psychology Machine Shop). Each chamber was equipped with a pellet dispenser that delivered 45-mg pellets into a recessed magazine when activated. Access to the magazine was detected by means of infrared detectors mounted across the opening of the recess. Two differently shaped panel lights were located on the right wall of the chamber above the magazine. The chambers contained a speaker connected to white noise and tone generators and a relay that delivered a 5kHz clicker stimulus. A computer equipped with GS3 software (Coulbourn Instruments, Allentown, PA) controlled the equipment and recorded the responses. Raw data were output and processed in Matlab (Mathworks, Natick, MA) to extract relevant response measures, which were analyzed in SPSS software (IBM analytics, Sydney, Aus).

Housing

Rats were housed singly and maintained on a 12 hour light-dark cycle, where all behavioral experiments took place during the light cycle. Rats had *ad libitum* access to food and water unless undergoing the behavioral experiment, during which they received sufficient chow to maintain them at ~85% of their free-feeding body weight. All experimental procedures were conducted in accordance with Institutional Animal Care and Use Committee of the US National Institute of Health guidelines.

General behavioral procedures

Trials consisted of 10s cues as described below. Trial types were interleaved in miniblocks, with the specific order unique to each rat and counterbalanced across groups. Intertrial intervals varied around a 6-min mean. Unless otherwise noted, daily training was divided into a morning (AM) and afternoon (PM) session. Inclusion of AM/PM as a factor in our analyses found no significant interactions with group so we collapsed this factor in the analyses in the main text.

Response measures

We measured entry into the food cup magazine to assess conditioned responding. Food cup entries were registered when the rat broke a photobeam placed across the opening of the food cup. This simple measure allows us to calculate a variety of metrics including response latency after cue onset, number of entries to the magazine during the cue, and the overall percentage of time spent in the food cup magazine during the cue. These metrics are generally correlated during conditioning, and all reflect to some extent the expectation of food delivery at the end of the cue in a task such as that used here^{51, 52}. One exception is in the relationship between number of entries and time spent in the food cup, which can vary depending on reward density in a task or behavioral setting^{52–56}. This tendency resulted in us focusing on two different metrics in the two experiments presented in this paper. Specifically, in Experiment 1, the overall rate of reward was relatively low, since only one cue was rewarded and it was only rewarded in a handful of trials in one phase of training. As a result, the rats exhibited relatively brief entries into the food cup, so number of entries was

the most reliable measure of conditioning. By contrast, in Experiment 2, the overall rate of reward was much higher, since two cues were rewarded, with two different rewards, in a larger number of trials. As a result, the rats spent much more time in the food cup each time they entered, so the percentage of time in the magazine was the most reliable measure of conditioning consistent with previous research^{52–56}. We note that in each case, the other measure yielded a qualitatively similar pattern of responding, however we have chosen to present the more reliable one for each experiment.

Histology and Immunohistochemistry

All rats were euthanized with an overdose of carbon dioxide and perfused with phosphate buffered saline (PBS) followed by 4% Paraformaldehyde (Santa Cruz Biotechnology Inc., CA). Fixed brains were cut in 40 μ m sections to examine fiber tip position under a 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 antisera (MAB318, 1: 600, EMD Millipore, Billerica, Massachusetts) followed by Alexa 568 secondary antisera (A11031, 1:1000, Invitrogen, Carlsbad, CA). Images of these brain slices were acquired by a fluorescence Virtual Slide microscope (Olympus America, NY) and later analyzed in Adobe Photoshop. We then counted the proportion of cells expressing eYFP that also co-stained for TH within the VTA of four subjects using sections taken from AP -5.0 mm to -6.0 mm. Positive staining was defined as a signal 2.5 times baseline intensity, with a cell diameter larger than 5 μ m, co-localized within cells reactive to DAPI staining. This encompasses the area likely to achieve good light penetration from our fibers.

Statistical Analyses

All statistics were conducted using SPSS 24 IBM statistics package. Generally, analyses were conducted using a mixed-design repeated-measures analysis of variance (ANOVA) with the exception of the data represented in Figure 3 which were conducted as one-way between-subjects ANOVA as appropriate. All analyses of simple main effects were planned and orthogonal and therefore did not necessitate controlling for multiple comparisons. Data distribution was assumed to be normal but homoscedasticity was not formally tested. Except for histological analysis, data collection and analyses were not performed blind to the conditions of the experiments.

Data Availability

The data that support the findings of this study, and any associated custom programs used for its acquisition, are available from the corresponding authors upon reasonable request.

Experiment 1

Subjects—Thirty-seven experimentally-naïve male ($n=19$) and female ($n=18$) Long-Evans transgenic rats of approximately 4 months of age at surgery and carrying a TH-dependent Cre expressing system (NIDA animals breeding facility) were used in this study. Sample sizes were chosen based on similar prior experiments which have elicited significant results with a similar number of rats. No formal power analyses were conducted. Rats were

randomly assigned to group and distributed equally by age, gender, and weight. Prior to data analysis, six rats were removed from the experiment due to illness, virus or cannula misplacement.

Blocking of Sensory Preconditioning—Training used a total of 6 different stimuli, drawn from stock equipment available from Coulbourn, and included 4 auditory (tone, siren, clicker, white noise) and two visual stimuli (flashing light, steady light). Assignment of these stimuli to the cues depicted in Fig 2 and described in the text was counterbalanced across rats within each modality (A, E were visual, and C, D, F and X were auditory).

Training began with 2 days of preconditioning. On the first day, the rats received 16 presentations of A→X, where a 10s presentation of A was immediately followed by a 10s presentation of X. On the second day, the rats received 8 presentations A→X alone, as well as 8 presentations each of three 10s compound cues (EF, AD, AC) followed by X (i.e. EF→X; AD→X; AC→X). On AC trials, light (473nm, 16–18mW output, Shanghai Laser & Optics Century Co., Ltd) was delivered into the VTA for 2s at a rate of 20Hz at the beginning of X; on AD trials, the same light pattern was delivered during the intertrial interval, 120–180s after termination of X. Following preconditioning, rats underwent 4 days of conditioning in which X was presented 24 times each day followed immediately by delivery of two 45-mg sucrose pellets (5TUT; TestDiet, MO). Finally, rats received a probe test in which each of the critical test cues (C, D, F) was presented 4 times alone and without reward.

Devaluation—A subset of the rats in the experimental ChR2 group ($n = 10$) underwent additional training after the probe test described above. These rats received reminder training in which X was again presented 12 times with reward, and then they were divided into two equal, performance-matched groups. Subsequently they received 30 minutes of access to 10 grams of the sucrose pellet reward to habituate them to receiving pellets outside of the training chamber, after which they began 3 days of training to devalue the sucrose pellet reward. Each day, one group (devalued; $n = 5$) received access to the sucrose pellets for 30 minutes, followed immediately by an intraperitoneal injection of a 0.15M solution of lithium chloride (LiCl; Sigma-Aldrich, MO) to induce nausea; the other group (non-devalued; $n = 5$) received the injections and were given a yoked amount of sucrose pellets approximately 6 hours later. Forty-eight hours following the third LiCl injection, all rats were given a final probe test where C was presented 12 times, alone and without reward, followed by a final consumption test in which all rats were received access to 10 grams of the sucrose pellets for 30 minutes.

Experiment 2

Subjects—Forty-one experimentally-naïve male ($n=33$) and female ($n=8$) Long-Evans transgenic rats of approximately 4 months of age at surgery and carrying a TH-dependent Cre expressing system (NIDA animal breeding facility) were used in this study. Sample sizes were chosen based on similar prior experiments which have elicited significant results with a similar number of rats. No formal power analyses were conducted. Rats were randomly

assigned to group and distributed equally by age, gender, and weight. Prior to data analysis, two rats were removed from the experiment due to illness, virus or cannula misplacement.

Sensory Preconditioning—Training used a total of 4 different auditory stimuli, drawn from stock equipment available from Coulbourn, which included tone, siren, clicker, and white noise. Assignment of these stimuli to the cues depicted in Fig. 5 and described in the text was counterbalanced across rats. Training began with one day of preconditioning where rats received 12 presentations of the A→X serial compound and 12 trials of the B→Y serial compound. Following preconditioning, rats began conditioning where they received 24 trials of X and 24 trials of Y each paired with a different reinforcer (either banana or grape pellets). Following 4 days of this training, rats received a probe test where cue A and B were each presented 6 times in the absence of any reinforcement.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by R01-MH098861 (YN) and by the Intramural Research Program at NIDA ZIA-DA000587 (GS). The authors 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 thank B Harvey and the NIDA Optogenetic and Transgenic Core and M Morales and the NIDA Histology Core for their assistance, and P Dayan and N Daw for their comments. The opinions expressed in this article are the authors' own and do not reflect the view of the NIH/DHHS.

References

- Schultz W. Dopamine neurons and their role in reward mechanisms. *Curr Opin Neurobiol.* 1997; 7:191–197. [PubMed: 9142754]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science.* 1997; 275:1593–1599. [PubMed: 9054347]
- Sutton RS, Barto AG. Toward a Modern Theory of Adaptive Networks - Expectation and Prediction. *Psychol Rev.* 1981; 88:135–170. [PubMed: 7291377]
- Nakahara H. Multiplexing signals in reinforcement learning with internal models and dopamine. *Current Opinion in Neurobiology.* 2014; 25:123–129. [PubMed: 24463329]
- Schultz W. Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews Neuroscience.* 2016; 17:183–195. [PubMed: 26865020]
- Tolman EC. Cognitive maps in rats and men. *Psychological Review.* 1948; 55:189–208. [PubMed: 18870876]
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci.* 2005; 8:1704–1711. [PubMed: 16286932]
- Glascher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron.* 2010; 66:585–595. [PubMed: 20510862]
- Colwill RM. An associative analysis of instrumental learning. *Current Directions in Psychological Science.* 1993; 2:111–116.
- Holland PC, Rescorla RA. The effects of two ways of devaluing the unconditioned stimulus after first and second-order appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes.* 1975; 1:355–363. [PubMed: 1202141]
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron.* 2011; 69:1204–1215. [PubMed: 21435563]

12. Steinberg EE, et al. A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*. 2013; 16:966–973. [PubMed: 23708143]
13. Eshel N, et al. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature*. 2015; 525:243–246. [PubMed: 26322583]
14. Chang CY, et al. Brief optogenetic inhibition of VTA dopamine neurons mimics the effects of endogenous negative prediction errors during Pavlovian over-expectation. *Nature Neuroscience*. 2016; 19:111–116. [PubMed: 26642092]
15. Tsai HC, et al. Phasic firing in dopamine neurons is sufficient for behavioral conditioning. *Science*. 2009; 324:1080–1084. [PubMed: 19389999]
16. Adamantidis AR, et al. Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *Journal of Neuroscience*. 2011; 31:10829–10835. [PubMed: 21795535]
17. 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]
18. 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]
19. Brogden WJ. Sensory pre-conditioning. *J Exp Psychol*. 1939; 25:323–332.
20. Blundell P, Hall G, Killcross S. Preserved sensitivity to outcome value after lesions of the basolateral amygdala. *Journal of Neuroscience*. 2003; 23:7702–7709. [PubMed: 12930810]
21. Jones JL, et al. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science*. 2012; 338:953–956. [PubMed: 23162000]
22. Rizley RC, Rescorla RA. Associations in Second-Order Conditioning and Sensory Preconditioning. *J Comp Physiol Psych*. 1972; 81:1-&.
23. Kamin, LJ. “Attention-like” processes in classical conditioning. In: Jones, MR., editor. *Miami Symposium on the Prediction of Behavior, 1967: Aversive Stimulation*. University of Miami Press; Coral Gables, Florida: 1968. p. 9-31.
24. Tobler PN, Dickinson A, Schultz W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *Journal of Neuroscience*. 2003; 23:10402–10410. [PubMed: 14614099]
25. 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]
26. Hollerman JR, Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*. 1998; 1:304–309. [PubMed: 10195164]
27. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*. 2012; 482:85–88. [PubMed: 22258508]
28. Takahashi YK, et al. The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for Learning from Unexpected Outcomes. *Neuron*. 2009; 62:269–280. [PubMed: 19409271]
29. Kakade S, Dayan P. Dopamine: generalization and bonuses. *Neural Networks*. 2002; 15:549–559. [PubMed: 12371511]
30. Horvitz JC, Stewart T, Jacobs B. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Research*. 1997; 759
31. Witten IB, et al. Recombinase-Driver Rat Lines: Tools, Techniques, and Optogenetic Application to Dopamine-Mediated Reinforcement. *Neuron*. 2011; 72:721–733. [PubMed: 22153370]
32. 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]
33. Parker NF, et al. Reward and choice encoding in terminals of midbrain dopamine neurons depends on striatal target. *Nature Neuroscience*. 2016; 19:845–854. [PubMed: 27110917]
34. 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]

35. Holland PC. Relations between Pavlovian-Instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*. 2004; 30:104–117. [PubMed: 15078120]
36. Dickinson A, Balleine BW. Motivational control of goal-directed action. *Animal Learning and Behavior*. 1994; 22:1–18.
37. Popescu AT, Zhou MR, Poo MM. Phasic dopamine release in the medial prefrontal cortex enhances stimulus discrimination. *Proceedings of the National Academy of Science*. 2016; 113:E3169–3176.
38. Mackintosh NJ. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol Rev*. 1975; 82:276–298.
39. Pearce JM, Hall G. A Model for Pavlovian Learning - Variations in the Effectiveness of Conditioned but Not of Unconditioned Stimuli. *Psychol Rev*. 1980; 87:532–552. [PubMed: 7443916]
40. 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 of London B: Biological Sciences*. 2011 rspb20110836.
41. Sadacca BF, Jones JL, Schoenbaum G. Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *eLife*. 2016; 5:e13665. [PubMed: 26949249]
42. Cone JJ, et al. Physiological state gates acquisition and expression of mesolimbic reward prediction signals. *Proceedings of the National Academy of Science*. 2016; 113:1943–1948.
43. Bromberg-Martin ES, Matsumoto M, Hong S, Hikosaka O. A pallidus-habenula-dopamine pathway signals inferred stimulus values. *Journal of Neurophysiology*. 2010; 104:1068–1076. [PubMed: 20538770]
44. Aitken TJ, Greenfield VY, Wassum KM. Nucleus accumbens core dopamine signaling tracks the need-based motivational value of food-paired cues. *Journal of Neurochemistry*. 2016 doi: 10.1111/jnc.13494.
45. Deserno L, et al. Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Science*. 2015; 112:1595–1600.
46. Eshel N, Tian J, Bukwich M, Uchida N. Dopamine neurons share common response function for reward prediction error. *Nature Neuroscience*. 2016; 19:479–486. [PubMed: 26854803]
47. Lammel S, et al. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron*. 2008; 57:760–773. [PubMed: 18341995]
48. Wimmer GE, Shohamy D. Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science*. 2012; 338:270–273. [PubMed: 23066083]
49. Robinson S, et al. Chemogenetic silencing of neurons in retrosplenial cortex disrupts sensory preconditioning. *Journal of Neuroscience*. 2014; 34:10982–10988. [PubMed: 25122898]
50. Johnson AW, Fenton AA, Kentros C, Redish AD. Looking for cognition in the structure within the noise. *Trends in Cognitive Sciences*. 2009; 13:55–64. [PubMed: 19135406]
51. Holland PC. Conditioned stimulus as a determinant of the form of Pavlovian conditioned response. *Journal of Experimental Psychology: Animal Behavior Processes*. 1977; 3:77–104. [PubMed: 845545]
52. McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. *Journal of Neuroscience*. 2011; 31:2700–2705. [PubMed: 21325538]
53. Holland PC, Gallagher M. Effects of amygdala central nucleus lesions on blocking and unblocking. *Behavioral neuroscience*. 1993; 107:235. [PubMed: 8484889]
54. Holland PC, Kenmuir C. Variations in unconditioned stimulus processing in unblocking. *Journal of Experimental Psychology: Animal Behavior Processes*. 2005; 31:155. [PubMed: 15839773]
55. Sharpe M, Killcross S. The prelimbic cortex contributes to the down-regulation of attention toward redundant cues. *Cerebral cortex*. 2014; 24:1066–1074. [PubMed: 23236210]
56. Burke KA, Franz TM, Miller DN, Schoenbaum G. The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature*. 2008; 454:340–344. [PubMed: 18563088]

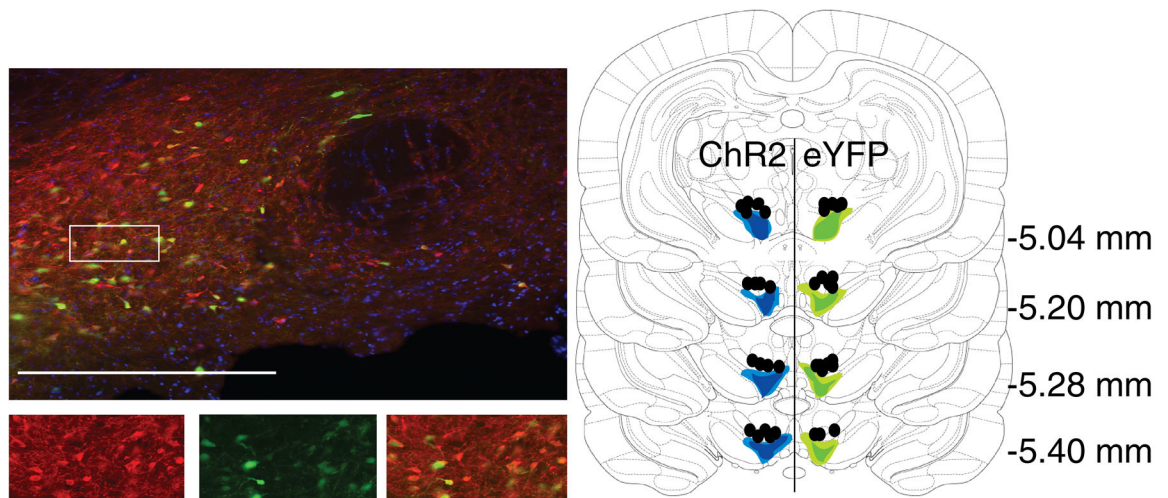


Figure 1. Immunohistochemical verification of Cre-dependent ChR2 and eYFP expression in tyrosine hydroxylase-expressing (TH+) neurons and fiber placements in the VTA
 Left images: 90% of YFP-expressing neurons (green) also expressed TH (red). Bottom images are an expansion of the region boxed at top. Right panel: Unilateral representation of the bilateral fiber placements and virus expression in each group. Fiber implants (black circles) were localized in the vicinity of eYFP (green) and ChR2 (blue) expression in VTA. Light shading represents the maximal and dark shading indicates the minimal spread of expression at each level. Scale bar = 20 μ m.

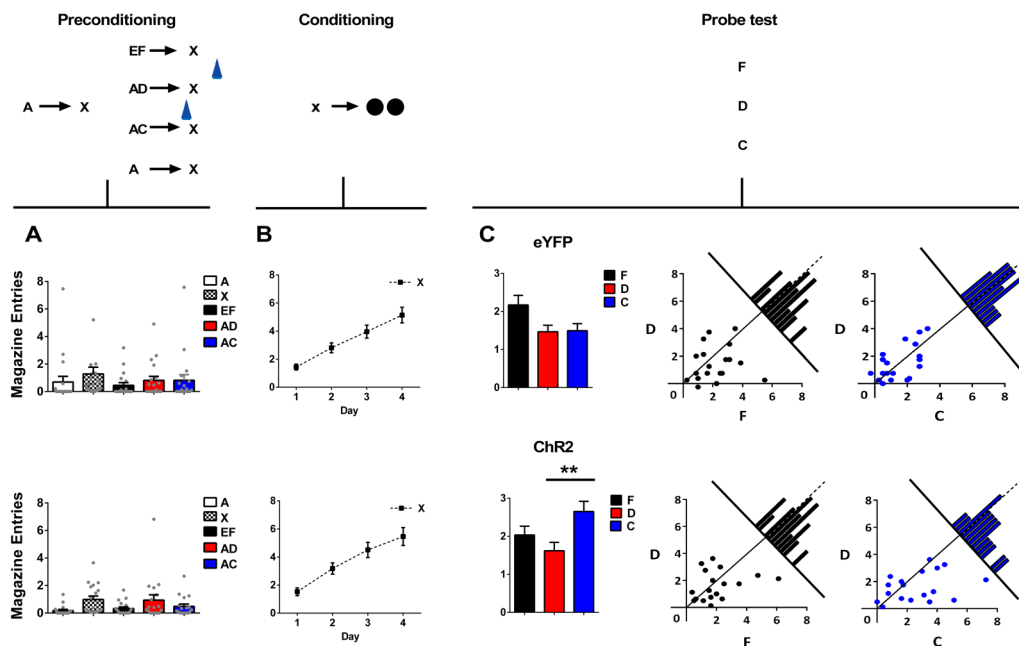


Figure 2. Brief optogenetic activation of VTA dopamine neurons strengthens associations between cues

Plots show number of magazine entries occurring during cue presentation across all phases of the blocking of sensory preconditioning task: preconditioning (A), conditioning (B) and the probe test (C). Probe test data are represented as the mean level of entries (left) or individual rats' responses to F and D (middle), or C and D (right). Top panel shows data from the eYFP control group ($n=19$), bottom panel shows data from the experimental Chr2 group ($n=18$). To the extent that responding to F and C are equal to D in scatterplots represented in panel C, points should congregate around the diagonal. Histograms along the diagonal reveal the frequency (subject counts) of difference scores in responding to the cues which fall within a particular range. VTA dopamine neurons were activated by light delivery at the beginning of presentations of X when preceded by audio-visual compound AC and during the ITI on AD trials, indicated by a blue symbol. A two-factor ANOVA on food cup entries during cue presentations (cue \times group) revealed no main effect ($F(4,140) = 1.52, p = 0.2$), nor any interaction with group ($F(4,140) = 0.276, p = 0.893$). A two-factor ANOVA (group \times day) on responding during conditioning (B) revealed a main effect of day ($F(3,105) = 39.71, p < 0.0001$) but neither main effect ($F(1,35) = 0.553, p = 0.46$) nor any interaction with group ($F(3,105) = 0.13, p = 0.94$). A two-factor ANOVA (cue \times group) on responding during presentation of cues F and D revealed a main effect of cue ($F(1,35) = 4.372, p = 0.044$) but no main effect ($F(1,35) = 0.001, p = 0.982$) or interaction with group ($F(1,35) = 0.287, p = 0.595$). A two-factor ANOVA (cue \times group) on responding to C and D revealed a main effect of cue ($F(1,35) = 4.599, p = 0.039$) and a significant interaction with group ($F(1,35) = 4.154, p = 0.049$). This interaction was due to a significant difference between responding to C and D in the Chr2 group ($F(1,35) = 8.52, p = 0.006$) but not in the eYFP group ($F(1,35) = 0.006, p = 0.940$). ** Indicates significance at $p < 0.01$. Error bars = SEM. Please see methods for comment on response measures, and Supplementary Information, Fig S4 for further details on responding during individual sessions in preconditioning.

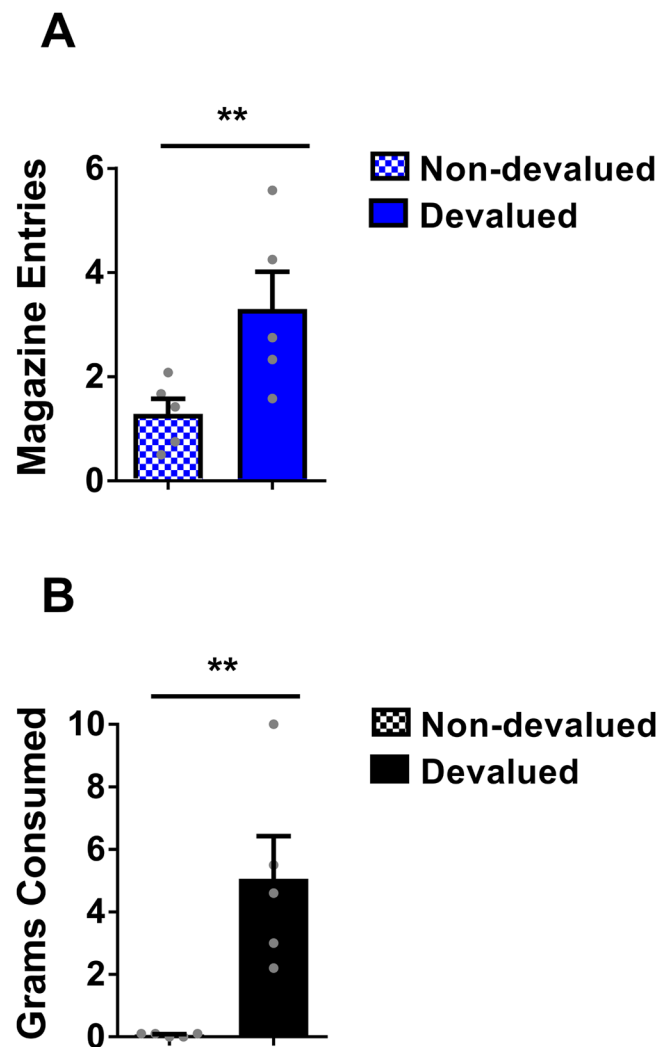


Figure 3. Conditioned responding resulting from learning supported by brief activation of VTA dopamine neurons is sensitive to devaluation of the predicted reward

(A) Magazine entries during presentation of C in the probe test following illness-induced devaluation of the predicted sucrose pellet reward. (B) Grams of sucrose pellets consumed in subsequent consumption test. A one-way ANOVA revealed a significant difference between responding to cue C ($F(1,8) = 6.777, p = 0.031$) and consumption of the sucrose food pellets ($F(1,8) = 13.425, p = 0.006$) in the devalued group relative to the non-devalued group. ** $p < 0.05$. Error bars = SEM.

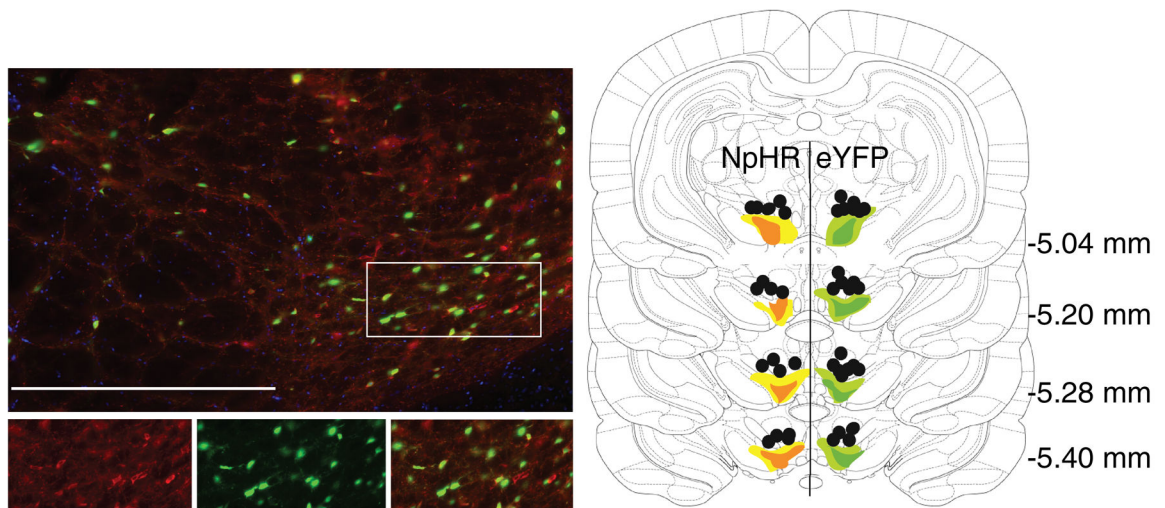


Figure 4. Immunohistochemical verification of Cre-dependent NpHR and eYFP expression in tyrosine hydroxylase-expressing (TH+) neurons and fiber placements in the VTA
 Left images: 90% of YFP-expressing neurons (green) also expressed TH (red). Bottom images are expansion of the region boxed at top. Right panel: Unilateral representation of the bilateral fiber placements and virus expression in each group. Fiber implants (black circles) were localized in the vicinity of eYFP (green) and NpHR (orange) expression in VTA. The light shading represents the maximal and the dark shading indicates the minimal spread of expression at each level. Scale bar = 20 μ m.

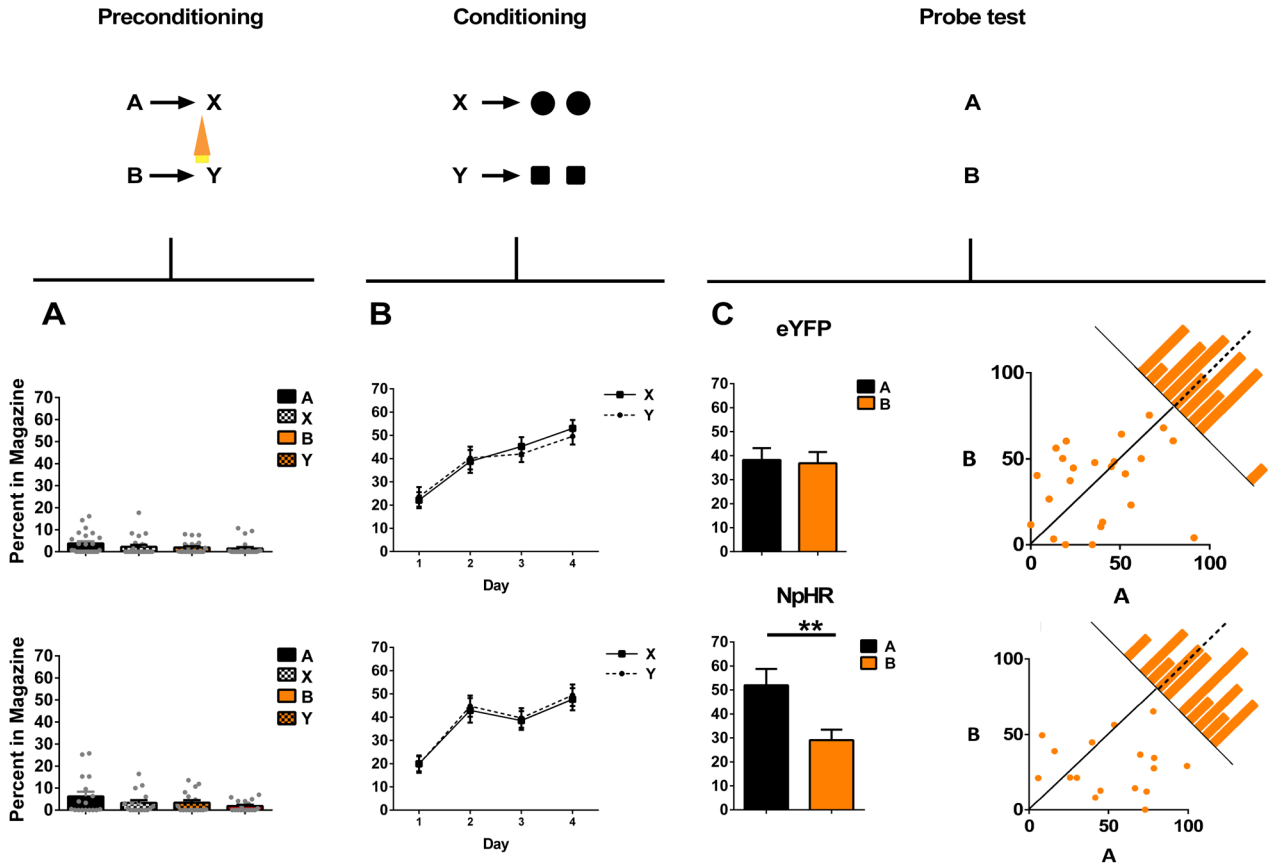


Figure 5. Brief optogenetic inhibition of dopamine neurons reduces the strength of associations between cues

Plots show the percentage of time spent in the magazine during cue presentation across all phases of the sensory preconditioning task: preconditioning (A), conditioning (B) and the probe test (C). Top panel shows data from the eYFP control group ($n=24$), bottom panel shows data from the experimental NpHR group ($n=17$). To the extent that responding to A is equal to B in scatterplots represented in panel C, points should congregate around the diagonal. Histograms along the diagonal reveal the frequency (subject counts) of difference scores in responding to the cues which fall within a particular range. VTA dopamine neurons were inhibited by light delivery in the 500ms before the offset of A and carried through the first 2s of X, indicated by an orange symbol. A two-factor ANOVA on food cup responding during cue presentations (cue \times group) in preconditioning (A) revealed no main effect ($F(1,39) = 1.88, p = 0.177$) nor any interaction with group ($F(3, 117) = 0.425, p = 0.736$). A three-factor ANOVA (cue \times group \times day) on data from conditioning (B) revealed a main effect of day ($F(3,105) = 43.181, p < 0.0001$) but no main effect of cue ($F(1,39) = 0.008, p = 0.927$), or group ($F(1,39) = 0.094, p = 0.761$) or any cue \times group interaction ($F(1,39) = 1.113, p = 0.298$). A two-factor ANOVA (cue \times group) revealed a main effect of cue ($F(1,39)=5.94, p=0.019$) and a significant cue by group interaction ($F(1,39)=4.68, p=0.037$). Subsequent comparisons showed that this interaction was due to a significant difference in responding to cues A and B in the NpHR group ($F(1,39) = 4.952, p = 0.012$) that was not present in the eYFP group ($F(1,39)=0.742, p = 0.483$). ** indicates significance at the

$p < 0.012$. Error bars = SEM. Please see methods for comment on response measures, and Supplementary Information, Fig S4 for further details on responding during individual sessions in preconditioning.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript