



Published in final edited form as:

Nat Neurosci. 2013 July ; 16(7): 966–973. doi:10.1038/nn.3413.

A Causal Link Between Prediction Errors, Dopamine Neurons and Learning

Elizabeth E. Steinberg^{1,2,*}, Ronald Keiflin^{1,*}, Josiah R. Boivin^{1,2}, Ilana B. Witten⁵, Karl Deisseroth⁶, and Patricia H. Janak^{1,2,3,4,*}

¹Ernest Gallo Clinic & Research Center, University of California, San Francisco, San Francisco, CA 94143, USA

²Graduate program in Neuroscience, University of California, San Francisco, San Francisco CA 94143, USA.

³Dept. of Neurology, University of California, San Francisco, San Francisco CA 94143, USA.

⁴Wheeler Center for the Neurobiology of Addiction, University of California, San Francisco, San Francisco CA 94143, USA.

⁵Princeton Neuroscience Institute & Department of Psychology, Princeton University, Princeton, NJ, 08544, USA.

⁶Department of Bioengineering, Department of Psychiatry and Behavioral Sciences, Howard Hughes Medical Institute, CNC Program, Stanford University, Stanford, CA 94305, USA.

Abstract

Situations where rewards are unexpectedly obtained or withheld represent opportunities for new learning. Often, this learning includes identifying cues that predict reward availability.

Unexpected rewards strongly activate midbrain dopamine neurons. This phasic signal is proposed to support learning about antecedent cues by signaling discrepancies between actual and expected outcomes, termed a reward prediction error. However, it is unknown whether dopamine neuron prediction error signaling and cue-reward learning are causally linked. To test this hypothesis, we manipulated dopamine neuron activity in rats in two behavioral procedures, associative blocking and extinction, that illustrate the essential function of prediction errors in learning. We observed that optogenetic activation of dopamine neurons concurrent with reward delivery, mimicking a prediction error, was sufficient to cause long-lasting increases in cue-elicited reward-seeking behavior. Our findings establish a causal role for temporally-precise dopamine neuron signaling in cue-reward learning, bridging a critical gap between experimental evidence and influential theoretical frameworks.

Users may view, print, copy, download and 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 to: pjanak@gallo.ucsf.edu.

* Denotes equal contribution.

Author contributions: E.E.S., R.K., and P.H.J. designed experiments. E.E.S., R.K., and J.R.B. performed experiments. I.B.W. and K.D. contributed reagents. E.E.S., R.K. and P.H.J. wrote the paper with comments from all authors.

INTRODUCTION

Much of the behavior of humans and other animals is directed towards seeking out rewards. Learning to identify environmental cues that provide information about where and when natural rewards can be obtained is an adaptive process that allows this behavior to be distributed efficiently. Theories of associative learning have long recognized that simply pairing a cue with reward is not sufficient for learning to occur. In addition to contiguity between two events, learning also requires the subject to detect a discrepancy between an expected reward and the reward that is actually obtained¹.

This discrepancy, or ‘reward prediction error’ (RPE), acts as a teaching signal used to correct inaccurate predictions. Presentation of unpredicted reward or reward that is better than expected generates a positive prediction error and strengthens cue-reward associations. Presentation of a perfectly predicted reward does not generate a prediction error and fails to support new learning. Conversely, omission of a predicted outcome generates a negative prediction error and leads to extinction of conditioned behavior. The error correction principle figures prominently in psychological and computational models of associative learning¹⁻⁶ but the neural bases of this influential concept remain to be definitively demonstrated.

In vivo electrophysiological recordings in non-human primates and rodents reveal that putative dopamine neurons in the ventral tegmental area (VTA) and the substantia nigra pars compacta respond to natural rewards such as palatable food⁷⁻⁹. Notably, the sign and magnitude of the dopamine neuron response is modulated by the degree to which the reward is expected. Surprising or unexpected rewards elicit strong increases in firing rate, while anticipated rewards produce little or no change^{8, 10, 11}. Conversely, when an expected reward fails to materialize, neural activity is depressed below baseline⁸⁻¹⁰. Reward-evoked dopamine release at terminal regions in vivo is also more pronounced when rewards are unexpected¹². On the basis of this parallel between RPE and dopamine responses, a current hypothesis suggests that dopamine neuron activity at the time of reward delivery acts as a teaching signal and causes learning about antecedent cues²⁻⁴. This conception is further supported by the observation that dopamine neurons are strongly activated by primary rewards before cue-reward associations are well-learned. As learning progresses and behavioral performance nears asymptote, the magnitude of dopamine neuron activation elicited by reward delivery progressively wanes^{7, 10}.

Though the correlative evidence linking reward-evoked dopamine neuron activity with learning is compelling, little causal evidence exists to support this hypothesis. Previous studies that attempted to address the role of prediction errors and phasic dopamine neuron activity in learning employed pharmacological tools, such as targeted inactivation of the VTA¹³, or administration of dopamine receptor antagonists¹⁴ or indirect agonists¹⁵. Such studies suffer from the major limitation that pharmacological agents alter the activity of neurons over long timescales and thus cannot determine the contribution of specific patterns of dopamine neuron activity to behavior. Genetic manipulations that chronically alter the actions of dopamine neurons by reducing or eliminating the ability of dopamine neurons to fire in bursts^{16, 17} do alter learning, but suffer from similar problems, as the impact of

dopamine neuron activity during specific behavioral events (such as reward delivery) cannot be evaluated. Other studies circumvented these issues using optogenetic tools that permit temporally-precise control of dopamine neuron activity; however these studies failed to utilize behavioral paradigms that explicitly manipulate reward expectation¹⁸⁻²¹, involve natural rewards^{20, 21}, or are suitable to assess cue-reward learning¹⁹. Thus, despite the prevalence and influence of the hypothesis that RPE signaling by dopamine neurons drives associative cue-reward learning, a direct link between the two has yet to be established.

To address this unresolved issue we capitalized on the ability to selectively control the activity of dopamine neurons in the awake, behaving rat with temporally-precise and neuron-specific optogenetic tools²¹⁻²³ in order to simulate naturally-occurring dopamine signals. We sought to determine whether activation of dopamine neurons in the VTA timed with the delivery of an expected reward would mimic a RPE and drive cue-reward learning using two distinct behavioral procedures.

First, we employed blocking, the associative phenomenon that best demonstrates the role of prediction errors in learning²⁴⁻²⁶. In a blocking procedure, the association between a cue and a reward is prevented (or 'blocked') if another cue present in the environment at the same time already reliably signals reward delivery²⁷. It is generally argued that the absence of an RPE, supposedly encoded by the reduced or absent phasic dopamine response to the reward, prevents further learning about the redundant cue^{4, 28}. We reasoned that artificial VTA dopamine neuron activation paired with reward delivery would mimic a positive prediction error and facilitate learning about the redundant cue. Next, we tested the role of dopamine neuron activation during extinction learning. Extinction refers to the observed decrease in conditioned responding that results from the reduction or omission of an expected reward. The negative prediction error, supposedly encoded by a pause in dopamine neuron firing, is proposed to induce extinction of behavioral responding^{4, 29}. We reasoned that artificial VTA dopamine neuron activation timed to coincide with the reduced or omitted reward would interfere with extinction learning. In both procedures, optogenetic activation of dopamine neurons at the time of expected reward delivery affected learning in a manner that was consistent with the hypothesis that dopamine neuron prediction error signaling drives associative learning.

RESULTS

Demonstration of associative blocking

The blocking procedure provides an illustration of the essential role of RPEs in associative learning. Consider two cues (e.g. a tone and a light) presented simultaneously (in compound) and followed by reward delivery. It has been shown that conditioning to one element of the compound is reduced (or 'blocked') if the other element has already been established as a reliable predictor of the reward²⁴⁻²⁷. In other words, despite consistent pairing between a cue and reward, the absence of a prediction error prevents learning about the redundant cue. In agreement with the conception that dopamine neurons encode prediction errors, putative dopamine neurons recorded in vivo exhibit little to no reward-evoked responses within a blocking procedure²⁸. The lack of dopamine neuron activity, combined with a failure to learn in the blocking procedure, is considered to be a key piece of

evidence (albeit correlative) linking dopamine RPE signals to learning. On the basis of this evidence, we determined that the blocking procedure would provide an ideal environment in which to test the hypothesis that RPE signaling by dopamine neurons can drive learning. According to this hypothesis, artificially activating dopamine neurons during reward delivery in the blocking condition when dopamine neurons normally do not fire should mimic a naturally-occurring prediction error signal and allow subjects to learn about the otherwise ‘blocked’ cue.

We first demonstrated associative blocking of reward-seeking (Fig. 1) using parameters suitable for subsequent optogenetic neural manipulation. Two groups of rats were initially trained to respond for a liquid sucrose reward (unconditioned stimulus, US) during an auditory cue in a “single cue” training phase. Subsequently, a combined auditory and visual cue was presented in a “compound” training phase and the identical sucrose US was delivered. For subjects assigned to the Blocking group, the same auditory cue was presented during single and compound phases, whereas distinct auditory cues were used for Control group subjects (Fig. 1a); in both phases, US delivery was contingent on the rat’s presence in the reward port during the cue (Fig. 1b). Hence, the critical difference between experimental groups is US predictability during the compound phase: because of its prior association with the previously-trained auditory cue, the US is expected for the Blocking group, whereas, for the Control group, its occurrence is unexpected. We measured conditioned responding as the amount of time spent in the reward port during the cue, normalized to an immediately preceding pre-cue period of equal length. Both groups showed equivalently high levels of conditioned behavior at the end of the single cue phase (two-way repeated-measures (RM) ANOVA, no effect of group or group \times day interaction, all p ’s >0.05), but differed in their performance when the compound cue was introduced (two-way RM ANOVA, main effect of group, $F_{1,21}=21.15$, $p<0.001$, and group \times day interaction, $F_{3,63}=11.63$, $p<0.001$), consistent with the fact that the association between the compound cue and US had to be learned by the Control group (Fig. 1c).

To determine whether learning about the visual cue introduced during compound training was affected by the predictability of reward, conditioned responding to unreinforced presentations of the visual cue alone was assessed one day later. Conditioned responding was reduced in the Blocking group as compared to Controls (Fig. 1 d-e; two-way RM ANOVA, main effect of group, $F_{1,21}=11.27$, $p=0.003$, no group \times trial interaction, $F_{2,42}=1.29$, $p=0.286$), indicating that new learning about preceding environmental cues occurs after unpredicted, but not predicted, reward in this procedure, in accord with prior findings^{28, 30}.

Reward-paired dopamine neuron activation drives learning

Putative dopamine neurons recorded in monkeys are strongly activated by unexpected reward, but fail to respond to the same reward if it is fully predicted^{10, 11}, including when delivered in a blocking condition²⁸. The close correspondence between dopamine neural activity and behavioral evidence of learning in this task suggests that positive RPEs caused by unexpected reward delivery activate dopamine neurons and lead to learning observed under control conditions. To test this hypothesis, we optogenetically activated VTA

dopamine neurons at the time of US delivery on compound trials in our blocking task to drive learning under conditions in which learning normally does not occur. We used parameters that we have previously established elicit robust, time-locked activation of dopamine neurons and neurotransmitter release in anesthetized animals or in vitro preparations²¹. We predicted that phasic dopamine neuron activation delivered coincidentally with fully-predicted reward would be sufficient to cause new learning about preceding cues.

Female transgenic rats expressing Cre recombinase under the control of the tyrosine hydroxylase (*Th*) promoter (*Th::Cre+* rats), or their wild-type littermates (*Th::Cre-* rats) were used to gain selective control of dopamine neuron activity as described previously²¹. *Cre+* and *Cre-* littermates received identical injections of a Cre-dependent virus expressing channelrhodopsin-2 (ChR2) in the VTA; chronic optical fiber implants were targeted dorsal to this region to allow for selective unilateral optogenetic dopamine neuron activation (Fig. S1, Fig. 2a). Three groups of rats were trained under conditions that normally result in blocked learning to the light cue (cue X; Fig. 2b). The behavioral performance of an experimental group (Paired*Cre+*), consisting of *Th::Cre+* rats receiving optical stimulation (1s train, 5ms pulse, 20 Hz) paired with the US during compound training (see Methods), was compared to the performance of two control groups that received identical training but differed either in genotype (Paired*Cre-*) or the time at which optical stimulation was delivered (Unpaired*Cre+*; optical stimulation during the intertrial interval; ITI) (Fig. 2c). Groups performed equivalently during single cue and compound training (Fig. 2d), suggesting that all rats learned the task and that the optical stimulation delivered during compound training did not disrupt ongoing behavior (two-way RM ANOVA revealed no significant effect of group or group \times day interaction; all p 's >0.111).

The critical comparison among groups occurred when the visual cue introduced during compound training was tested alone in an unreinforced session. Paired*Cre+* subjects responded more strongly to the visual cue on the first test trial than subjects from either control group (Fig. 2e-f), indicating greater learning. A two-way RM ANOVA revealed a significant interaction between group and trial ($F_{4,50}=3.819$, $p=0.009$) and a trend towards a main effect of group ($F_{2,25}=3.272$, $p=0.055$). Planned post-hoc comparisons showed a significant difference between the Paired*Cre+* group and Paired*Cre-* ($p=0.005$) or Unpaired*Cre+* ($p<0.001$) controls on the first test trial, while control groups did not differ (Unpaired*Cre+* vs. Paired*Cre-* $p=0.155$; Fig. 2e-f). This result demonstrates that unilateral VTA dopamine neuron activation at the time of US delivery is sufficient to cause new learning about preceding environmental cues. The observed dopamine neuron-induced learning enhancement was temporally specific, as responding to the visual cue was blocked in the Unpaired*Cre+* group receiving optical stimulation outside of the cue and US periods. Importantly, Paired*Cre+* and Unpaired*Cre+* rats received equivalent stimulation and this stimulation was equally reinforcing (Fig. S2a-c), so discrepancies in the efficacy of optical stimulation between *Cre+* groups cannot explain the observed behavioral differences.

One possible explanation for the behavioral changes we observed in the blocking experiment is that optical stimulation of dopamine neurons during compound training served to increase the value of the paired sucrose reward. Such an increase in value would result in a RPE (although not encoded by dopamine neurons) and unblock learning. We found

however that the manipulation of dopamine neuron activity during the consumption of one of two equally-preferred, distinctly-flavored sucrose solutions did not change the relative value of these rewards, measured as reward preference; Fig. S3,S4 and Methods). This suggests that the unblocked learning about the newly-added cue \times was not the result of increased reward value induced by manipulating dopamine neuron activity.

Dopamine neuron activation slows extinction

Negative prediction errors also drive learned behavioral changes. For example, after a cue-reward association has been learned, decrementing or omitting the expected reward results in decreased reward-seeking behavior. Dopamine neurons show a characteristic pause in firing in response to reward decrements or omissions⁸⁻¹⁰, and this pause is proposed to contribute to decreased behavioral responding to cues after reward decrement^{4, 29}. Having established that optogenetically activating dopamine neurons can drive new learning about cues under conditions in which dopamine neurons normally do not change their firing patterns from baseline levels, we next tested whether similar artificial activation at a time when dopamine neurons normally decrease firing could counter decrements in behavioral performance produced by US value reductions. *Th::Cre+* and *Th::Cre-* rats that received unilateral ChR2-containing virus infusions and optical fiber implants targeted to the VTA (Fig. S1), were trained to respond for sucrose whose availability was predicted by an auditory cue. One day after the last training session, the auditory cue was presented but water was substituted for the sucrose US (Downshift test; Fig. 3a). PairedCre+ and PairedCre- rats received dopamine neuron optical stimulation (3s train, 5ms pulse, 20 Hz) concurrent with water delivery when they entered the reward port during the cue; UnpairedCre+ rats received stimulation during the ITI. One day later, rats were given a downshift recall session identical to the first downshift test except that no optical stimulation was given. The purpose of the recall session was to determine if optical stimulation had caused long-lasting behavioral changes. Cue responding was measured as the percent time spent in the reward port during the cue normalized to a pre-cue baseline (Fig. 3b-d), and as the latency to enter the reward port after cue onset (Fig. 3e-g).

As seen in Fig. 3b and 3e, all groups acquired the initial cue-US association; a two-way RM ANOVA revealed no significant effects of group or group \times day interactions at the end of training (all p 's > 0.277). During the downshift test, PairedCre- and UnpairedCre+ group performance rapidly deteriorated. This was evident on a trial-by-trial basis (Fig. 3c,f, left) and when cue responding was averaged across the entire downshift test session (Fig. 3c,f, right). In contrast, PairedCre+ rats receiving optical stimulation concurrent with water delivery showed much reduced (Fig. 3c) or no (Fig. 3f) decrement in behavioral responding. Two-way RM ANOVAs revealed significant effects of group and group \times trial interactions for both time spent in the port during the cue ($F_{2,28}=11.12$, $p<0.001$ and $F_{18,252}=1.953$, $p=0.013$) and latency to respond after cue onset ($F_{2,28}=12.463$, $p<0.001$ and $F_{18,252}=4.394$, $p<0.001$). Planned post-hoc comparisons demonstrated that PairedCre+ rats differed significantly from controls in both time and latency ($p<0.001$) while control groups did not differ from each other ($p>0.375$). Notably, some group differences persisted into the downshift recall session wherein no stimulation was delivered (latency: main effect of group, $F_{2,28}=4.597$, $p=0.019$; Fig. 3g). These data indicate that phasic VTA dopamine

neuron activation can partially counteract performance decrements associated with reducing reward value.

We next determined if our optical manipulation would be effective if the expected reinforcer was omitted entirely (Fig. 4a). Rats used in the Downshift experiment (see Methods) were trained on a new cue-reward association. All rats learned the new association (Fig. 4b, e); a two-way RM ANOVA revealed no significant effects of group or group \times day interactions at the end of training (all p 's > 0.242). Subsequently all rats were given an extinction test, during which the expected sucrose reward was withheld. Instead, PairedCre⁺ and PairedCre⁻ rats received optical stimulation (3s train, 5ms pulse, 20 Hz) of dopamine neurons at the time of expected US delivery, while UnpairedCre⁺ rats received optical stimulation during the ITI. One day later, rats were given an extinction recall session where neither the US nor optical stimulation were delivered to determine if prior optical stimulation caused long-lasting behavioral changes.

During the extinction test, PairedCre⁺ rats spent more time in the reward port during the cue and responded to the cue more quickly as compared to both PairedCre⁻ and UnpairedCre⁺ rats (Fig. 4c,f); two-way RM ANOVAs revealed significant effects of group and/or group \times trial interactions for both measures (percent time: $F_{2,28}=40.054$, $p<0.001$ and $F_{18,252}=0.419$, $p=0.983$; latency: $F_{2,28}=3.827$, $p=0.034$ and $F_{18,252}=2.047$, $p=0.008$), and these behavioral differences persisted into the extinction recall session (Fig. 4d,g; two-way RM ANOVAs, significant main effects of group and group \times trial interactions ($F>2$, $p<0.01$ in all cases). Hence, VTA dopamine neuron activation at the time of expected reward is sufficient to sustain conditioned behavioral responding when expected reward is omitted. For both reward downshift and omission, the behavioral effects of dopamine neuron stimulation were temporally specific, as UnpairedCre⁺ rats responded less than PairedCre⁺ rats despite receiving more stimulation during the test sessions (Fig. S2d,g), and despite verification that this stimulation is equally reinforcing in both Cre⁺ groups (Fig. S2e,f & h,i).

Despite causing significant behavioral changes during extinction, optogenetic activation of dopamine neurons failed to maintain reward-seeking behavior at pre-extinction levels. This may be due to the inability of our dopamine neuron stimulation to fully counter the expected decrease in dopamine neuron firing during reward omission or downshift. Alternatively, this may reflect competition between the artificially-imposed dopamine signal and other neural circuits specialized to inhibit conditioned responding when this behavior is no longer advantageous, as has been previously proposed^{31, 32}.

Estrus cycle can modulate dopaminergic transmission under some circumstances³³. Notably, although female rats were used in these studies, we tracked estrus stage during a behavioral session where dopamine neurons were stimulated and failed to observe correlations between estrus cycle stage and behavioral performance (Fig. S5).

DISCUSSION

Here we demonstrate that reward prediction error signaling by dopamine neurons is causally related to cue-reward learning. We leveraged the temporal precision afforded by optogenetic tools to mimic endogenous RPE signaling in VTA dopamine neurons in order to determine how these artificial signals impact subsequent behavior. Using an associative blocking procedure, we observed that increasing dopamine neuron activity during reward delivery could drive new learning about antecedent cues that would not normally guide behavior. Using extinction procedures, we observed that reductions in conditioned responding that normally accompany decreases in reward value are attenuated when dopamine neuron activity is increased at the time of expected reward. Importantly, the behavioral changes we observed in all experiments were long-lasting, persisting twenty-four hours after dopamine neurons were optogenetically activated, and temporally-specific, failing to occur if dopamine neurons were activated at times outside of the reward consumption period. Taken together, our results demonstrate that RPE signaling by dopamine neurons is sufficient to support new cue-reward learning and modify previously-learned cue-reward associations.

Our results clearly establish that artificially activating VTA dopamine neurons at the time that a natural reward is delivered (or expected) supports cue-elicited responding. A question of fundamental importance is why this occurs. In particular, for the blocking study, one possibility is that dopamine stimulation acted as an independent reward, discriminable from the paired sucrose reward, which initiated the formation of a parallel association between the reward-predictive cue and dopamine stimulation itself. However this explanation assumes cue independence and would require the rat to compute two simultaneous yet separate prediction errors controlling, respectively, the strength of two separate associations (cue A \rightarrow sucrose; cue X \rightarrow dopamine stimulation). Indeed, the assumption of cue independence was challenged¹ specifically because separate prediction errors cannot account for the phenomenon of blocking. If each cue generated its own independent prediction error, then the preconditioning of one cue would not affect the future conditioning of other cues; but it does, as the blocking procedure shows. Blocking demonstrates that cues presented simultaneously interact and compete for associative strength. Hence, it is unlikely that a parallel association formed between reward-predictive cues and dopamine stimulation can account for our results. Of interest, putative dopamine neurons do not appear to encode a sensory representation of reward since they do not discriminate among rewards based on their sensory properties²⁹, thus it is not obvious how dopamine neuron activation at the same time as natural reward delivery could be perceived as distinct from that reward.

Although previous studies suggest otherwise^{34, 35}, another related possibility is that optical activation of dopamine neurons induces behavioral changes by directly enhancing the value of the paired natural reward. To address this possibility, we conducted a control study based on the conception that high-value rewards are preferred over less valuable alternatives. We paired dopamine stimulation with consumption of a flavored, and thus discriminable, sucrose solution; we reasoned that if dopamine stimulation served to increase the value of a paired reward, this should manifest as an increased preference for the stimulation-paired reward over a distinctly-flavored but otherwise identical sucrose solution. However, we observed that reward that was previously paired with dopamine stimulation was preferred

equivalently to one that was not. This result does not support the interpretation that optical dopamine stimulation supported learning in our experiments by increasing the value of the sucrose reward.

Alternatively, the behavioral changes we observed in PairedCre+ rats could reflect the development of a conditioned place preference for the location where optical stimulation was delivered (i.e., the reward port), as has been previously demonstrated²⁰. If this were the case, we should have observed generalized increases in reward-seeking behavior across the entire test session. Critically, our primary behavioral metric (time spent in the reward port during the cue) was normalized to pre-cue baseline levels. If optical stimulation had induced non-specific increases in reward-seeking behavior, our normalized measure should have approached zero. However, we found that reward-seeking was specifically elevated during cue presentation. While we observed robust group differences in our normalized measures, a separate analysis of the absolute percent time spent in the port in the pre-cue baseline period during any test session revealed no significant group differences in baseline responding (all p 's > 0.17). Together, these findings indicate that the behavioral changes we observed are unlikely to be the result of a conditioned place preference.

Instead, the most parsimonious explanation for our results is that dopamine stimulation reproduced a RPE. Theories of associative learning hold that simple pairing, or contiguity, between a stimulus and reward or punishment is not sufficient for conditioning to occur; learning requires the subject to detect a discrepancy or 'prediction error' between the expected and actual outcome which serves to correct future predictions¹. Although compelling correlative evidence indicated that dopamine neurons are well-suited to provide such a teaching signal, remarkably little proof existed to support this notion. For this reason, our results represent an advance over previous work. While prior studies that also utilized optogenetic tools to permit temporally-precise control of dopamine neuron activity demonstrated that dopamine neuron activation is reinforcing, these studies did not establish the means by which this stimulation can reinforce behavior. Because we used behavioral procedures in which learning is driven by reward prediction errors, our data establish the critical behavioral mechanism (RPE) through which phasic dopamine signals timed with reward cause learning.

Through which cellular and circuit mechanisms could this dopamine signal cause learning to take place? Though few in number, VTA dopamine neurons send extensive projections to a variety of cortical and subcortical areas and are thus well-positioned to influence neuronal computation^{2, 36-38}. Increases in dopamine neuron firing during unexpected reward could function as a teaching signal used broadly within efferent targets to strengthen neural representations that facilitate reward receipt^{39, 40}, possibly via alterations in the strength and direction of synaptic plasticity^{37, 41-43}. Because our artificial manipulation of dopamine neuron activity produced behavioral changes that lasted at least twenty-four hours after the stimulation ended, such dopamine-induced, downstream changes in synaptic function may have occurred; additionally, both natural cue-reward learning⁴⁴ and optogenetic stimulation of dopamine neurons⁴⁵ alter glutamatergic synaptic strength onto dopamine neurons themselves, providing another possible basis for the long-lasting effects of dopamine neuron activation on behavior. One or both of these synaptic mechanisms may underlie the

behavioral changes reported here. While the physiological consequences of optogenetic dopamine neuron activation have been investigated in in vitro preparations and in anesthetized rats, to fully explore these synaptic mechanisms, a first critical step is to define the effects of optical activation on neuronal firing and on dopamine release in the awake behaving subject.

Here we focused on the role of dopamine neuron activation at the time of reward. Another hallmark feature of dopamine neuron firing during associative learning is the gradual transfer of neural activation from reward delivery to cue onset. Early in learning when cue-reward associations are weak, dopamine neurons respond robustly to the occurrence of reward and weakly to reward-predictive cues. As learning progresses neural responses to the cue become more pronounced and reward responses diminish¹⁰. While our current results support the idea that reward-evoked dopamine neuron activity drives conditioned behavioral responding to cues, the function(s) of cue-evoked dopamine neuron activity remain a fruitful avenue for investigation.

Possible answers to this question have already been proposed. This transfer of the dopamine teaching signal from the primary reinforcer to the preceding cue is predicted by temporal-difference models of learning⁴⁶. In such models, the back-propagation of the teaching signal allows the earliest predictor of the reward to be identified, thereby delineating the chain of events leading to reward delivery^{2, 6, 46}. Alternatively, or in addition, cue-evoked dopamine may encode the cue's incentive value, endowing the cue itself with some of the motivational properties originally elicited by the reward, thereby making the cue desirable in its own right³⁴. Using behavioral procedures that allow a cue's predictive and incentive properties to be assessed separately, a recent study provided evidence for dopamine's role in the acquisition of cue-reward learning for the latter, but not the former process⁴⁷. Such behavioral procedures could also prove useful to determine in greater detail how learning induced by mimicking RPE signals impacts cue-induced conditioned responding. These and other future attempts to define the precise behavioral consequences of dopamine neuron activity during cues and rewards will further refine our conceptions of the role of dopamine RPE signals in associative learning.

ONLINE METHODS

Subjects and surgery

115 female transgenic rats (Long-Evans background) were used in these studies; 68 rats expressed Cre Recombinase under the control of the *tyrosine hydroxylase* promoter (*Th::Cre* +), and 47 rats were their wild-type littermates (*Th::Cre*-). All rats weighed >225g at the time of surgery. During testing (except the flavor preference study), rats were mildly food-restricted to 18g of lab chow per day given after the conclusion of daily behavioral sessions; on average rats maintained >95% free-feeding weight. Water was available *ad libitum* in the home cage. Rats were singly-housed under a 12:12 light/dark cycle, with lights on at 7am. The majority of behavioral experiments were conducted during the light cycle. Animal care and all experimental procedures were in accordance with guidelines from the National Institutes of Health and were approved in advance by the Gallo Center Institutional Animal Care and Use Committee. Stereotaxic surgical procedures were used for VTA infusion of

Cre-dependent virus (Efl α -DIO-ChR2-eYFP;²⁰) and optical fiber placement as in²¹, with the exception that DV coordinates were adjusted to account for the smaller size of female rats as follows: DV -8.1 and -7.1 mm below skull surface for virus infusions, and -7.1 mm for optical fiber implants.

Behavioral procedures

All behavioral experiments were conducted >2 weeks post-surgery; sessions that included optical stimulation were conducted >4 weeks post-surgery.

Apparatus—Behavioral sessions were conducted in sound-attenuated conditioning chambers (Med Associates Inc.). The left and right walls were fitted with reward delivery ports; computer-controlled syringe pumps located outside of the sound-attenuating cubicle delivered sucrose solution or water to these ports. The left wall had two nosepoke ports flanking the central reward delivery port; each nosepoke port had three LED lights at the rear. Chambers were outfitted with 2700 Hz pure tone and whitenoise auditory stimuli, both delivered at 70 dB, as well as a 28V chamber light above the left reward port. During behavioral sessions, the pure tone was “pulsed” at 3 Hz (0.1s on/0.2s off) to create a stimulus easily distinguished from continuous whitenoise.

Reward delivery—All experiments (except the flavor preference study) involved delivery of a liquid sucrose solution (15% w/v) during the presentation of auditory or combined auditory-visual cues. During each cue, entry into the active port triggered a 3s delivery of sucrose solution (0.1 mL). After a 2s timeout, another entry into the port (or the rat’s continued presence at the port) triggered an additional 3s reward delivery. This 5s cycle could be repeated up to 6 times per 30s trial, depending on the rat’s behavior. For sessions where optical stimulation was delivered, the laser was activated each time sucrose was delivered (or expected) as illustrated in Fig. 2C, 3A, 4A. This method of reward delivery, where reward and optical stimulation were both contingent on the rat’s presence in the active port, was used for all experiments as it allowed for the coincident delivery of natural rewards and optical stimulation and maximized the temporal precision of reward expectation.

Blocking procedure—Rats received a one-day habituation session where all auditory and visual cues used during future training sessions, as well as the sucrose reward, were presented individually (3 presentations of each cue, 5min ITI; ~ 60 reward deliveries, 1min ITI). This session was intended to minimize unconditioned responses to novel stimuli and shape reward-seeking behavior to the correct (left) reward port. Next, rats underwent single-cue training where one of two auditory cues (whitenoise or pulsed-tone, counterbalanced across subjects) was presented for 30s on a variable interval (VI) 4min schedule for 10 trials per session. Sucrose was delivered during each cue as described above. After 14–15 sessions of single cue training, compound cue training commenced and lasted for 4 days. During this phase either the same auditory cue used in single cue training (Blocking groups) or a new auditory cue (Control group) was presented simultaneously with a visual cue. The visual cue consisted of the chamber light, which was the sole source of chamber illumination, flashing on/off at 0.3 Hz (1 sec on, 2 sec off). Sucrose reward was delivered as described above

during this phase. A probe test was administered 24 hours after the conclusion of the last compound training session to assess conditioned responding to the visual cue. During this session the visual cue was presented alone in the absence of sucrose, auditory cues or optical stimulation.

Downshift procedure—Rats received one session where sucrose reward was delivered to the active (right) port (50 deliveries, VI-30 sec) to shape reward-seeking behavior to this location. Subsequently, rats were trained to respond for sucrose during an auditory cue (whitenoise) as described above in 11 daily sessions. 24 hours later, a downshift test session was administered that was identical to previous training sessions except that water was substituted for sucrose and optical stimulation was delivered coincidentally. 24 hours later, a downshift recall test was given, in which water was delivered during the cue, but optical stimulation did not occur.

Extinction procedure—This experiment was conducted two weeks after the end of the Downshift experiment with the same subjects; group assignment for Cre+ rats was shuffled between experiments. Rats received one session of sucrose reward delivery to the opposite (left) port used in the downshift test to shape reward-seeking behavior to this location. Subsequently rats were trained to respond for sucrose during an auditory cue (pulsed-tone) as described above in 6 daily training sessions. 24 hours later, an extinction test session was administered identical to previous training sessions except that no reward was given and optical stimulation was delivered at the time that the sucrose reward had been available in previous training. 24 hours later, an extinction recall test was administered in which the auditory cue was presented but no reward or optical stimulation was delivered.

Intracranial self-stimulation (ICSS)—Upon completion of the experiments described above, all rats were given 4 daily 1-hr sessions of ICSS training, as described previously²¹. Food restriction ceased at least 24 hours before the first ICSS session. A response at the nosepoke port designated as “active” resulted in the delivery of a train of light pulses matched to the stimulation parameters used in that subject’s previous behavioral experiment (1s/20Hz for rats in blocking or flavor preference studies, 3s/20Hz for rats in downshift or extinction studies).

Flavor preference study

Pre-training and pre-exposure: Rats were initially trained to drink unflavored 15% (w/v) sucrose solution from the reward port in the conditioning chambers (0.1 ml delivered on VI-30s schedule; 50 deliveries). Rats were then given overnight access to 40 ml each of the flavored sucrose solutions (15% sucrose (w/v) + 0.15% Kool-aid tropical punch or grape flavors (w/v)) in their home cage to ensure that all subjects had sampled both flavors before critical consumption tests. Home cages were equipped with two bottle slots; prior to the start of the experiment both slots were occupied by water bottles to reduce possible side bias.

Home cage consumption tests: Water bottles were removed from the home cage 15–30 min prior to the start of consumption tests. A standardized procedure was used to ensure that rats briefly sampled both flavors before free access to the solutions began. The purpose of

this procedure was to make sure that rats were aware that both flavors were available, so that any measured preference reflected true choice. The experimenter placed a flavor bottle on the left side of the cage until the rat consumed the solution for 2–3s. This bottle was removed, and a second bottle containing the other flavored solution was placed on the right side of the cage until the rat consumed the new solution for 2–3s. The second bottle was then removed, and both bottles were simultaneously placed on the home cage to start the test. 10 minutes later, bottles were removed and amounts consumed were recorded. The cage side assigned to each flavor (left or right) alternated between consumption tests to control for possible side bias.

Flavor training with optical stimulation: 24 hours after the home cage baseline consumption test, daily flavor training began in the conditioning chambers (8 sessions total). Only one flavored sucrose solution was available per day; training days with each flavor were interleaved. One of the two flavors was randomly assigned for each rat to be the “stimulated” flavor. On training days when the “stimulated” flavor was available, optical stimulation was either paired with reward consumption for PairedCre+ and PairedCre– groups, or explicitly unpaired (presented during the ITI at times when no reward was available) for UnpairedCre+ rats (Fig. S4). Flavored sucrose was delivered to a reward port on a VI-30s schedule, with the exception that each reward had to be consumed before the next would be delivered. A reward was considered to be consumed if the rat maintained presence in the port for 1s or longer. Sessions lasted until the maximum of 50 rewards were consumed or one hour elapsed, whichever occurred first. The final preference test was conducted 24 hours after the last flavor training session.

Optical activation

Intracranial light delivery in behaving rats was achieved as described²¹. For all experiments, 5ms light pulses were delivered with a 50ms inter-pulse interval (i.e., 20 Hz). For blocking and flavor preference experiments, 20 pulses were used (1s of stimulation). For downshift and extinction experiments, 60 pulses were used (3 s). Data from sessions where light output was compromised because of broken or disconnected optical cables was discarded and these subjects were excluded from the study. This criterion led to the exclusion of 1 rat from each of the Blocking and Extinction experiments, and 4 rats from the self-stimulation protocol.

Assessment of estrus cycle

Stage of estrus cycle was assessed by vaginal cytological examination using well-established methods⁴⁸. After behavioral sessions (Downshift study), the tip of a moistened cotton swab was gently inserted into the exterior portion of the vaginal canal and then rotated to dislodge cells from the vaginal wall. The swab was immediately rolled onto a glass slide, and the sample preserved with spray fixative (Spray-Cyte, Fisher Scientific) without allowing the cells to dry. Samples were collected over five consecutive days to ensure observation of multiple estrus cycle stages. This was done to improve the accuracy of determining estrus cycle stage on any single day of the experiment. Slides were then stained with a modified Papanicolaou staining procedure as follows: 50% ethyl alcohol, 3min; tap water, 10 dips (x2); Gill’s hematoxylin 1, 6min; tap water, 10 dips (x2); Scott’s water, 4min; tap water, 10 dips (x2); 95% ethyl alcohol, 10 dips (x2); modified orange-greenish 6 (OG-6), 1min; 95% ethyl

alcohol, 10 dips; 95% ethyl alcohol 8 dips; 95% ethyl alcohol 6 dips; modified eosin azure 36 (EA-36), 20min; 95% ethyl alcohol, 40 dips; 95% ethyl alcohol, 30 dips; 95% ethyl alcohol, 20 dips; 100% ethyl alcohol, 10 dips (x2); xylene, 10 dips (x2); coverslip immediately. All staining solutions (Gill's hematoxylin 1, OG-6, and EA-36) were sourced from Richard Allen Scientific. Estrus cycle stage was determined by identifying cellular morphology characteristic to each phase according to previously described criteria⁴⁸.

Histology

Immunohistochemical detection of YFP and TH was performed as described previously²¹. Although optical fiber placements and virus expression varied slightly from subject to subject, no subject was excluded based on histology (Fig. S1).

Data analysis

Counterbalancing procedures were used to form experimental groups balanced in terms of age, weight, conditioning chamber used, cue identity and behavioral performance in the sessions preceding the experimental intervention. Conditioned responding was measured as the amount of time spent in the reward port during cue presentation, normalized by subtracting the time spent in the port during a pre-cue period of equal length. Note that during reinforced training sessions, this measure is not a pure index of learning since the time spent in the port during the cue also reflects time spent consuming sucrose. For the blocking experiment, we focused exclusively on this measure because it proved to be particularly robust. Critically, during the blocking test itself, this measure *is* a pure index of learning because no reward is delivered during this session. For other experiments we also measured the latency to enter the reward port after cue onset. Pilot experiments and power analyses for both the blocking and the extinction study indicated that 8-10 subjects per group allowed for detection of differences between experimental and control conditions, with $\alpha=.05$ and $\beta=.80$. In cases where behavioral data from individual subjects varied from the group mean by more than two standard deviations (calculated with data from all subjects included), these subjects were excluded as statistical outliers (2 rats from the blocking experiment and 3 each from the downshift and extinction experiments) and their data were not further analyzed. Behavioral measures were analyzed using a mixed factorial ANOVA with the between-subjects factor of experimental group and the within-subjects factor of session or trial, followed by planned Student Newman-Keul's tests when indicated by significant main effects or interactions. For all tests, $\alpha=0.05$, and all statistical tests were 2-sided. By their design, the experiments focused on three planned comparisons (PairedCre+ vs UnpairedCre+; PairedCre+ vs PairedCre-; UnpairedCre+ vs PairedCre-). We found no major deviation from the assumptions of the ANOVA. For the cases where normality or equal variance was questionable, the results of the ANOVA were confirmed by non-parametric tests (Kruskal-Wallis followed by post-hoc Dunn's test). While explicit blinding procedures were not employed, experimental group allocation was not noted on subject cage cards, and all behavioral data were collected automatically via computer. Blocking and extinction experiments are replications of pilot experiments. Though based on a pilot study, the flavor preference study was conducted as described only one time, but with sufficient sample size to make statistical inferences.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful for the technical assistance of M. Olsman, L. Sahuque and R. Reese, and for critical feedback from H. Fields during manuscript preparation. This research was supported by US National Institutes of Health grants DA015096, AA17072 and funds from the State of California for medical research on alcohol and substance abuse through UCSF (PHJ); an NSF Graduate Research Fellowship (EES); a National Institutes of Health New Innovator award (IBW); and US National Institutes of Health grants from NIMH and NIDA, the Michael J Fox Foundation, the Howard Hughes Medical Institute, and the Defense Advanced Research Projects Agency (DARPA) Reorganization and Plasticity to Accelerate Injury Recovery (REPAIR) Program (KD).

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.(1972)
2. Glimcher PW. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc Natl Acad Sci U S A*. 2011; 108(Suppl 3):15647–54. [PubMed: 21389268]
3. Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci*. 1996; 16:1936–47. [PubMed: 8774460]
4. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997; 275:1593–9. [PubMed: 9054347]
5. Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci*. 2000; 23:473–500. [PubMed: 10845072]
6. Sutton RS, Barto AG. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev*. 1981; 88:135–70. [PubMed: 7291377]
7. Schultz W, Apicella P, Ljungberg T. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J Neurosci*. 1993; 13:900–13. [PubMed: 8441015]
8. 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; 482:85–8. [PubMed: 22258508]
9. Roesch MR, Calu DJ, Schoenbaum G. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*. 2007; 10:1615–24. [PubMed: 18026098]
10. Hollerman JR, Schultz W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci*. 1998; 1:304–9. [PubMed: 10195164]
11. Mirenowicz J, Schultz W. Importance of unpredictability for reward responses in primate dopamine neurons. *J Neurophysiol*. 1994; 72:1024–7. [PubMed: 7983508]
12. Day JJ, Roitman MF, Wightman RM, Carelli RM. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci*. 2007; 10:1020–8. [PubMed: 17603481]
13. Takahashi YK, et al. The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*. 2009; 62:269–80. [PubMed: 19409271]
14. Iordanova MD, Westbrook RF, Killcross AS. Dopamine activity in the nucleus accumbens modulates blocking in fear conditioning. *Eur J Neurosci*. 2006; 24:3265–70. [PubMed: 17156387]
15. O’Tuathaigh CM, et al. The effect of amphetamine on Kamin blocking and overshadowing. *Behav Pharmacol*. 2003; 14:315–22. [PubMed: 12838037]
16. Parker JG, et al. Absence of NMDA receptors in dopamine neurons attenuates dopamine release but not conditioned approach during Pavlovian conditioning. *Proc Natl Acad Sci U S A*. 2010; 107:13491–6. [PubMed: 20616081]

17. Zweifel LS, et al. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proc Natl Acad Sci U S A*. 2009; 106:7281–8. [PubMed: 19342487]
18. Adamantidis AR, et al. Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J Neurosci*. 2011; 31:10829–35. [PubMed: 21795535]
19. Domingos AI, et al. Leptin regulates the reward value of nutrient. *Nat Neurosci*. 2011; 14:1562–8. [PubMed: 22081158]
20. Tsai HC, et al. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 2009; 324:1080–4. [PubMed: 19389999]
21. Witten IB, et al. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*. 2011; 72:721–33. [PubMed: 22153370]
22. Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat Neurosci*. 2005; 8:1263–8. [PubMed: 16116447]
23. Zhang F, Wang LP, Boyden ES, Deisseroth K. Channelrhodopsin-2 and optical control of excitable cells. *Nat Methods*. 2006; 3:785–92. [PubMed: 16990810]
24. Kamin, LJ. Attention-like. In: Jones, MR., editor. *processes in classical conditioning*. University of Miami Press; Coral Gables, FL: 1968.
25. Kamin, LJ. Selective association and conditioning. Honig, N. J. M. a. W. K., editor. *Dalhousie University Press*; Halifax, NS: 1969.
26. Kamin, LJ. Predictability, surprise, attention and conditioning. Church, B. A. C. a. R. M., editor. *Appleton-Century-Crofts*; New York, NY: 1969.
27. Holland PC. Unblocking in Pavlovian appetitive conditioning. *J Exp Psychol Anim Behav Process*. 1984; 10:476–97. [PubMed: 6491608]
28. Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature*. 2001; 412:43–8. [PubMed: 11452299]
29. Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol*. 1998; 80:1–27. [PubMed: 9658025]
30. 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; 454:340–4. [PubMed: 18563088]
31. Daw ND, Kakade S, Dayan P. Opponent interactions between serotonin and dopamine. *Neural Netw*. 2002; 15:603–16. [PubMed: 12371515]
32. Peters J, Kalivas PW, Quirk GJ. Extinction circuits for fear and addiction overlap in prefrontal cortex. *Learn Mem*. 2009; 16:279–88. [PubMed: 19380710]
33. Becker JB. Gender differences in dopaminergic function in striatum and nucleus accumbens. *Pharmacol Biochem Behav*. 1999; 64:803–12. [PubMed: 10593204]
34. Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev*. 1998; 28:309–69. [PubMed: 9858756]
35. Wassum KM, Ostlund SB, Balleine BW, Maidment NT. Differential dependence of Pavlovian incentive motivation and instrumental incentive learning processes on dopamine signaling. *Learn Mem*. 2011; 18:475–83. [PubMed: 21693635]
36. Beckstead RM, Domesick VB, Nauta WJ. Efferent connections of the substantia nigra and ventral tegmental area in the rat. *Brain Res*. 1979; 175:191–217. [PubMed: 314832]
37. Fields HL, Hjelmstad GO, Margolis EB, Nicola SM. Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. *Annu Rev Neurosci*. 2007; 30:289–316. [PubMed: 17376009]
38. Swanson LW. The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Res Bull*. 1982; 9:321–53. [PubMed: 6816390]
39. Reynolds JN, Wickens JR. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw*. 2002; 15:507–21. [PubMed: 12371508]
40. Wickens JR, Horvitz JC, Costa RM, Killcross S. Dopaminergic mechanisms in actions and habits. *J Neurosci*. 2007; 27:8181–3. [PubMed: 17670964]

41. Gerfen CR, Surmeier DJ. Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci.* 2011; 34:441–66. [PubMed: 21469956]
42. Reynolds JN, Hyland BI, Wickens JR. A cellular mechanism of reward-related learning. *Nature.* 2001; 413:67–70. [PubMed: 11544526]
43. Tye KM, et al. Methylphenidate facilitates learning-induced amygdala plasticity. *Nat Neurosci.* 2010; 13:475–81. [PubMed: 20208527]
44. Stuber GD, et al. Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons. *Science.* 2008; 321:1690–2. [PubMed: 18802002]
45. Brown MT, et al. Drug-driven AMPA receptor redistribution mimicked by selective dopamine neuron stimulation. *PLoS One.* 2010; 5:e15870. [PubMed: 21209835]
46. Suri RE. TD models of reward predictive responses in dopamine neurons. *Neural Netw.* 2002; 15:523–33. [PubMed: 12371509]
47. Flagel SB, et al. A selective role for dopamine in stimulus-reward learning. *Nature.* 2010; 469:53–7. [PubMed: 21150898]

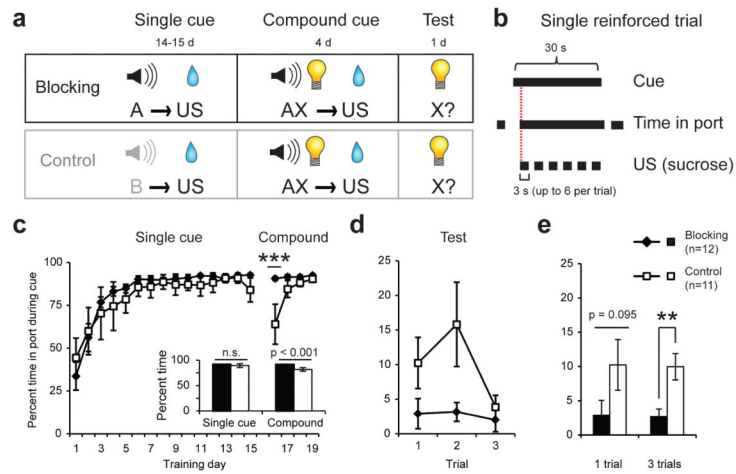


Fig. 1. Behavioral demonstration of the blocking effect

(a) Experimental design of the blocking task. (b) During reinforced trials, sucrose delivery was contingent upon reward port entry during the 30s cue. After entry, sucrose was delivered for 3s followed by a 2s timeout. Up to 6 sucrose rewards could be earned per trial, depending on the rats' behavior. (c) Performance across all single cue and compound training sessions. Inset, mean performance among groups over the last four days of single-cue training did not differ; controls showed reduced behavior during compound training (** $p < 0.001$). (d) Performance during visual cue test. The blocking group exhibited reduced responding to the cue at test relative to controls (main effect of group, $p = 0.003$, group \times trial interaction, $p = 0.286$). (e) Visual cue test performance for the first trial and the average of all three trials. The blocking group showed reduced cue responding for the 3-trial measure (** $p = 0.003$) but were not different on the first trial ($p = 0.095$). For all figures, values depicted are means and error bars represent SEM.

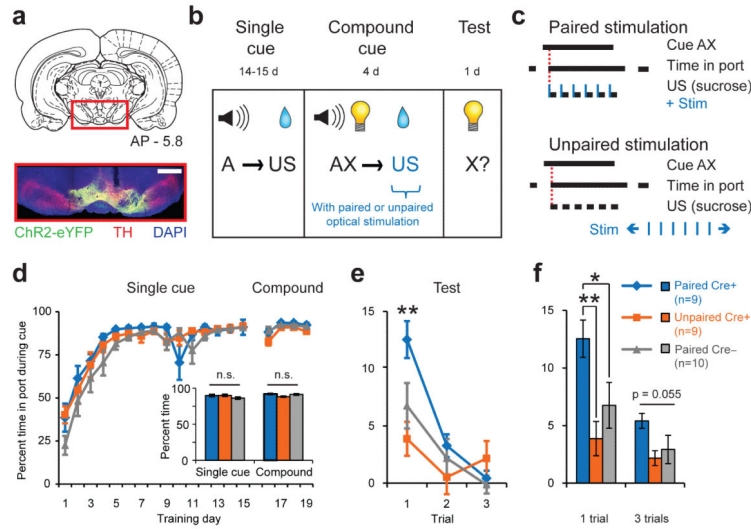


Fig. 2. Dopamine neuron stimulation drives new learning

(a) Example histology from a *Th::Cre*⁺ rat injected with a Cre-dependent ChR2-containing virus. Vertical track indicates optical fiber placement above VTA. Scale bar = 1mm. (b) Experimental design for blocking task with optogenetics. All groups received identical behavioral training according to the “blocking” group design in Fig. 1a. (c) Optical stimulation (1s train, 5ms pulse, 20 Hz, 473nm) was synchronized with sucrose delivery in Paired (Cre⁺ and Cre⁻), but not Unpaired (Cre⁺), groups. (d) Performance across all single cue and compound training sessions. Inset, no group differences over last four days of single cue training or during compound training. (e) Performance during visual cue test. The PairedCre⁺ group exhibited increased responding to the cue relative to both control groups at test on the first trial (***p*<0.005). (f) Visual cue test performance for the first trial and all three trials averaged. The PairedCre⁺ group exhibited increased cue responding relative to controls for the 1-trial measure (PairedCre⁺ vs. UnpairedCre⁺, ***p*=0.005, PairedCre⁺ vs. PairedCre⁻, **p*=0.025, PairedCre⁻ vs. UnpairedCre⁺, *p*=0.26); there was a trend for a group effect for the 3-trial average (main effect of group, *p*=0.055).

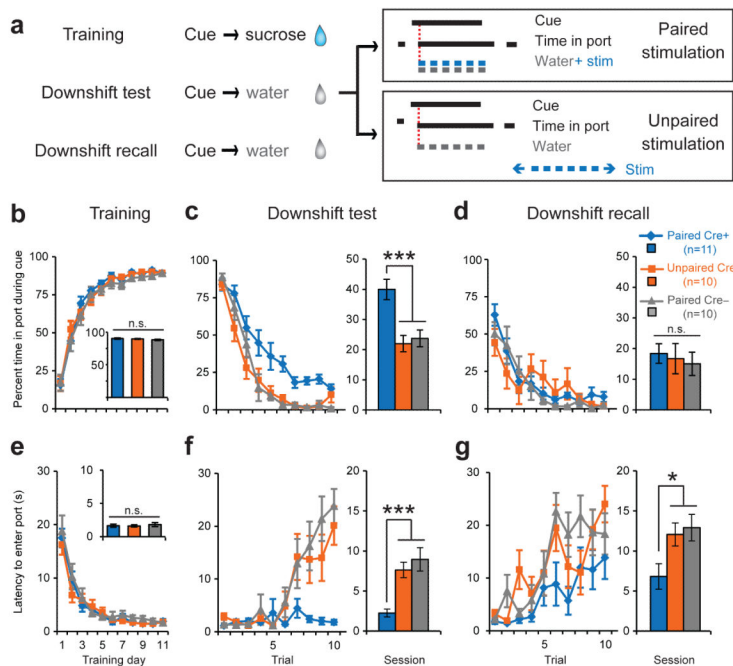


Fig. 3. Dopamine neuron stimulation attenuates behavioral decrements associated with a downshift in reward value

(a) Experimental design for reward downshift test. Optical stimulation (3s train, 5ms pulse, 20 Hz, 473nm) was either paired with the water “reward” (PairedCre+ and Cre- groups) or explicitly unpaired (UnpairedCre+). (b) Percent time in port during the cue across training sessions. Inset, no difference in average performance during the last two training sessions. (c) Percent time in port during the cue for the downshift test. Data are displayed for single trials (left) and as a session average (right). PairedCre+ rats exhibited increased time in port compared to controls (PairedCre+ vs. UnpairedCre+, *** $p < 0.001$, PairedCre+ vs. PairedCre-, *** $p < 0.001$, PairedCre- vs. UnpairedCre+, $p = 0.691$). (d) Percent time in port during the cue for downshift recall. Data are displayed for single trials (left) and as a session average (right). There were no group differences during this phase (2-way RM ANOVA, main effect of group $p = 0.835$). (e) Latency to enter the reward port after cue onset. Inset, no group differences during last two training sessions. (f) As in C, but for latency. PairedCre+ rats responded faster to the cue compared to controls during the downshift test (PairedCre+ vs. UnpairedCre+, *** $p < 0.001$, PairedCre+ vs. PairedCre-, *** $p < 0.001$, PairedCre- vs. UnpairedCre+, $p = 0.375$). (g) As in D, but for latency. PairedCre+ rats responded faster to the cue compared to controls during downshift recall (PairedCre+ vs. UnpairedCre+, * $p = 0.024$, PairedCre+ vs. PairedCre-, * $p = 0.025$, PairedCre- vs. UnpairedCre+, $p = 0.706$).

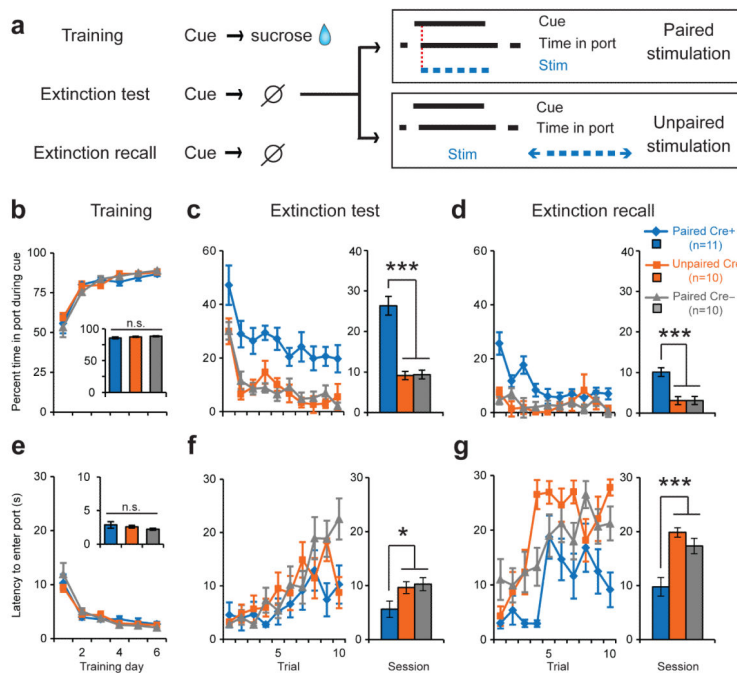


Fig. 4. Dopamine neuron stimulation attenuates behavioral decrements associated with reward omission

(a) Experimental design for extinction test. Note that the same subjects from the downshift experiment were used for this procedure, with Cre+ groups shuffled between experiments (see Methods). Optical stimulation (3s train, 5ms pulse, 20 Hz, 473nm) was delivered at the time of expected reward for Paired groups and during ITI for UnpairedCre+ rats. (b) Percent time in port during the cue across training sessions. Inset, no difference in average performance during the last two training sessions. (c) Percent time in port during the cue for the extinction test. Data are displayed for single trials (left) and as a session average (right). PairedCre+ rats exhibited increased time in port compared to controls (PairedCre+ vs. UnpairedCre+, *** $p < 0.001$, PairedCre+ vs. PairedCre-, *** $p < 0.001$, PairedCre- vs. UnpairedCre+, $p = 0.920$). (d). Percent time in port during the cue for extinction recall. Data are displayed for single trials (left) and as a session average (right). PairedCre+ rats exhibited increased time in port compared to controls (PairedCre+ vs. UnpairedCre+, *** $p < 0.001$, PairedCre+ vs. PairedCre-, *** $p < 0.001$, PairedCre- vs. UnpairedCre+, $p = 0.984$). (e) Latency to enter the reward port after cue onset. Inset, no group differences during last two training sessions. (f) As in C, but for latency. PairedCre+ rats responded faster to the cue compared to controls during the extinction test (PairedCre+ vs. UnpairedCre+, * $p = 0.038$, PairedCre+ vs. PairedCre-, * $p = 0.04$, PairedCre- vs. UnpairedCre+ $p = 0.727$). (g) As in D, but for latency. PairedCre+ rats responded faster to the cue compared to controls during extinction recall (PairedCre+ vs. UnpairedCre+, *** $p < 0.001$, PairedCre+ vs. PairedCre-, *** $p < 0.001$, PairedCre- vs. UnpairedCre+, $p = 0.211$).